Repeated Contests with Private Information*

Greg Kubitz†

August 22, 2017

Abstract

In repeated contests with private information, weak contestants prefer to appear strong while strong contestants prefer to appear weak. In contrast to a single contest, this leads to an equilibrium where effort is not strictly monotonic in ability and allows for a less able contestant to win against a contestant of higher ability. While the aggregate payoffs of contestants are higher per contest than in the single contest benchmark, aggregate output per contest is lower. Depending on the economic setting, the presence of private information can lead to productive or allocational inefficiencies.

1 Introduction

Contests are frequently used to stimulate effort from economic agents. These contests are often dynamic and offer multiple prizes, as in the case of tournaments and employee competitions. In both of these settings, there is extensive literature discussing how to best design contests to maximize the output of the contestants.[1] However, the behavior of economic agents in repeated contests is not fully characterized in situations where agents have private information about productive ability or value of winning the tournament.[2]

In this paper, we study repeated contests in a framework designed to capture both moral hazard (hidden effort choice) and adverse selection (privately known abilities). That is, when contestants' abilities are private information, the contestants must consider the signaling effect that exerting effort in early contests will have in future contests. Contrary to the conventional wisdom that all contestants want to appear strong to their opponents, countervailing incentives emerge in this setting and strategies are not strictly monotone. These incentives can create multiple inefficiencies. First, overall output in repeated contests is lower relative to the single contest benchmark. In the contests described above, this reduction in output is a negative welfare result. Second,

*I would like to thank Tomasz Sadzik, Marek Pycia, Moritz Meyer-ter-Vehn, Simon Board, John Asker, Bill Zame, Ichiro Obara, Simon Loertscher, Kyle Woodward, Robert Ziethammer, Sergio Parreiras and in particular Jernej Copic for his guidance throughout this project. I am also grateful to seminar audiences at Melbourne and UNC Chapel Hill for their numerous helpful comments. All remaining errors are attributable to me.

†NYU Stern, gkubitz@stern.nyu.edu

[1] Fullerton and McAfee (1999) and Ye (2007) explain how multi-stage all pay auctions can be optimal in cases of costly selection of participants or costly entry for participants. For labor market competitions, see Lazear and Rosen (1981) and Rosen (1986) and more recently Ridlon and Shin (2013), Ederer (2010), and Aoyagi (2010).

[2] A model of repeated all-pay auctions where bidders have private valuation is contained within the current model.

a contestant with low ability may beat a contestant with high ability in the first contest or both contests. In multi-stage tournaments, this may prevent the best contestant from winning the tournament, or even making the later rounds.

We consider the simplest setting that captures the signaling incentives of repeated contests: two contestants, who have either low or high ability, competing in two successive contests. In each contest, the contestants exert effort with the goal of producing the most output. The player who does so wins a prize.[3] The amount of effort it takes to produce output depends on individual ability, which is privately known by each contestant. After the first contest, the output of each contestant is publicly observed and players can update their beliefs about their opponents' ability. Given this additional information, contestants choose a new level of effort for the second contest. The contestants choose their effort levels to maximize their total payoff over the two contests.

We show there is a unique symmetric equilibrium for this repeated contest game. The equilibrium strategies of both high and low ability contestants reflect the trade-off between success in the first contest and optimal positioning for success in the second contest. The complementarity of ability and effort would lead to a high level of output from high ability contestants and a low level of output from low ability contestants if there was only a single contest. However, entering the second contest, a contestant with high ability will always prefer to have his opponent believe they have low ability. Likewise, a contestant with low ability wants to appear to have high ability. Concern about the outcomes in both contests leads to an equilibrium that has partial pooling in the first contest, i.e. there are outputs which can be produced by either low ability or high ability contestants. Low ability players who produce output in this range are *bluffing* while high ability players who do so are *sandbagging*.[4]

Both tactics are used in the first contest with the purpose of depressing the effort of opponents in the second contest. In particular, contestants care primarily about the effort choice of opponents who have similar ability. For example, a low ability opponent will be discouraged when they observe high output in the first contest, and will lead them to put less effort into the second contest. Therefore, low ability contestants have an incentive to put a high level of effort into the first contest. On the other hand, a high ability opponent would increase their effort in the second contest if they observe a high output in the first contest. In order to avoid this escalation, high ability contestants will exert relatively little effort in the first contest.

The partial pooling of outputs in the first contest allows for a high ability contestant to lose to a low ability contestant even when the two contestants are ex-ante symmetric. In a single contest, contestants use separating strategies, and a high ability contestant will always win against a low ability contestant when they are ex-ante symmetric. Additionally, when the cost of production is represented by a monomial, the equilibrium strategies of the repeated contest have a lower expected output per contest than the expected output of equilibrium strategies of a single contest. While bluffing in

[3]For a general description of static (one-shot) games of this kind, see Siegel (2009).

[4]The terms *sandbag* and *bluff* are used in the literature to describe a player signaling to his opponent that he is weak when he is actually strong and strong when he is actually weak, respectively. These terms originate from the game of poker. In poker, *sandbagging* is when a player calls or does not increase the pot when he believes he has the better hand. *Bluffing* is when a player bids up the pot when he does not think he has the best hand.

the first contest increases expected output of low ability contestants, sandbagging by high ability contestants has the opposite effect on output. Further, information revealed about the ability of contestants after the first contest often creates asymmetries that decrease the competitiveness of the second contest. This will reduce the total output in the second contest.[5]

Uniqueness of equilibrium in dynamic games with signaling is not common and stems from the countervailing incentives in the first contest. To derive this equilibrium, we first use the construction in Siegel (2014) to find the unique equilibrium of the second contest subgame for any set of abilities and beliefs that emerge from the first contest. The incentives to sandbag and bluff clearly emerge from the equilibrium payoffs in the second contest. For high ability contestants, expected payoffs strictly decrease in the other contestant's belief about their ability. For low ability contestants, these payoffs strictly increase in the opponent's belief. The desire of each type of contestant to appear as the other type not only leads to a unique partial pooling equilibrium in the first contest, it also rules out any additional equilibria. In signaling games, undesirable off-equilibrium path beliefs can often be used to construct additional perfect Bayesian equilibria. In this game, however, there are no beliefs that are undesirable to both high and low ability contestants. Therefore any belief used in an off-path output of the first contest will cause a deviation that unravels the potential equilibrium.

Our work contributes to the extensive literature investigating the manipulation of information in dynamic competitions. In a majority of signaling games, competitors are incentivized to bluff in order to discourage opponents. This behavior is observed in labor competitions as over working before a midterm evaluation (Ederer (2010)), in dynamic auctions as jump bidding (Avery (1998)) and in duopoly competition as excess production when firm costs are uncertain (Mirman et al. (1993) and Bonatti et al. (2016)). Sandbagging, as described in Rosen (1986), is often used to lull opponents into a false sense of security. This manipulation also appears as bid-shading in repeated first price auctions, see Ortega Reichert (2000) and Bergemann and Horner (2014).

In our setting, the incentives for strong competitors to sandbag and weak competitors to bluff is due to (i) the existence of private information, (ii) the partial revelation of this information through actions during the competition and (iii) the all-pay nature of the contest. This two directional distortion of strategies is also identified in Hörner and Sahuguet (2007) where bidders are able to signal their value through a fixed jump bid prior to an all-pay auction. Bidders with moderate values will sometimes use this jump bid while bidders with high values may not.[6] Our model suggests that the non-monotonic strategies in this setting are not an artifact of the limited strategy space before the auction.

Repeated contests with private information are also studied in Münster (2009), where contestants may or may not value winning the contest. This leads to an equilib-

[5]The reduction of output from asymmetry is also identified in Baye et al. (1993), Che and Gale (2003) and Terwiesch and Xu (2008). It differs from the discouragement effect caused by falling behind in early stages of a contest of complete information and uncertain outcomes as in Harris and Vickers (1987). Because each period is a separate contest in our model, reduced output is not due to this discouragement effect. See Konrad (2012) for a detailed survey of dynamic contests under complete information.

[6]Konrad and Morath (2017) also identify these incentives for contestants who can decide to form an alliance ahead of a conflict.

rium where contestants who do value the contest will sometimes not participate in the first contest. In the current paper, the more general set up of two levels of ability also identifies the incentive of low ability contestants to bluff.

Contests that proceed in two stages, such as elimination tournaments, reduce the number of participants in the second round. When the goal of the contest designer is to maximize the highest effort in the contest, Moldovanu and Sela (2006) show that a two-stage contest design where winners of the two first-round contests face each other in a second round is preferable to a static contest between all contestants. Similarly, Fullerton and McAfee (1999) argue that in many cases restricting a research contest to two contestants through a contestant selection auction is cheapest method of procurement. Both results are reliant on efficient entry into the contest, i.e. the best two contestants enter the second round. In these models efficient entry is achieved, as either outputs from the first round are not observed or private information is not assumed to persist to the second round. Our results show that if outputs or bids in the first round are publicly observable, top contestants would be worried about revealing information about their value before the last round. This leads to a non-separating equilibrium in the first round of a tournament which causes inefficient entry into the second round. Non-existence of a separating equilibrium in the first round of a four player, two round tournament has previously been shown by Zhang and Wang (2009) in the all-pay auction setting with ex-ante symmetric contestants who have values of winning the tournament distributed on a compact support. Additionally, our analysis shows that even in the case where the best contestants enter the second round, the incentive to prevent bid escalation will lead to reduced output in the first round and lead to a less competitive second round.

The rest of the paper is organized as follows. In section 2 we formalize the repeated contest model. In sections 3 and 4 we characterize the equilibrium of this model by backwards induction, focusing on the second contest in section 3 and the first contest in section 4. In section 5 we discuss welfare implications of the equilibrium and section 6 concludes.

2 Model

Two ex-ante identical contestants are independently endowed with ability, a^i, for $i = 1, 2$. Each contestant is equally likely to have low ability, $a = a_\ell$, or high ability, $a = a_h$. Ability is normalized so that $a_\ell = 1$ and $a_h > 1$, and the endowment of ability is private information for each contestant. After the initial draw of types, the abilities of the contestants are fully persistent.

There are two sequential contests in which the contestants compete by choosing effort, e, and producing output, x. Effort and ability are complimentary, and the output function takes the form $x(a, e) = a \cdot e$. At the end of the first contest, the outputs of each contestant from the first contest, (x_1^1, x_1^2), become public information. Contestants use this information to update their beliefs about their opponent's ability prior to competing in the second contest. This belief is denoted by $\mu^{-i}(x_1^i)$ and is the probability that contestant i has high ability given the information available to contestant $-i$

after the first contest. Because these outputs are commonly observed, contestants first order beliefs are sufficient for characterizing each contestant's information in the second contest.

In each contest, the contestant that produces the most output receives a prize. If the two contestants produce the same output, then the prize is given randomly, each contestant winning with equal probability. Prizes have the same value across contests and contestants, which we normalize to one. Both contestants bear the cost of their effort in each contest regardless of the its outcome. This cost, $c(e)$, is the same for high and low ability contestants. The cost function is assumed to be twice differentiable on the non-negative reals, increasing and weakly convex, with the cost of zero effort being zero. The payoff of contestant i given the ability and effort choice of each contestant are

$$\tilde{\pi}^i(a^i, e^i, a^{-i}, e^{-i}) = \begin{cases} 1 - c(e^i), & x(a^{-i}, e^{-i}) < x(a^i, e^i) \\ 1/2 - c(e^i), & x(a^{-i}, e^{-i}) = x(a^i, e^i) \\ -c(e^i), & x(a^{-i}, e^{-i}) > x(a^i, e^i) \end{cases}$$

Given an output of their opponent, the expected payoffs of each contestant is equal to the probability that the contestant wins less his cost of effort. Here we abuse notation and let $x^i = x(a^i, e^i)$ for $i = 1, 2$.

$$E[\tilde{\pi}^i(a^i, e^i)] = \Pr(x^{-i} < x(a^i, e^i)) + \frac{1}{2}\Pr(x^{-i} = x(a^i, e^i)) - c(e^i).$$

Since contestants know their own ability and the relationship between effort and output is deterministic, choosing effort level is equivalent to choosing output.[7] Therefore, strategies are written in terms of output to ease comparisons of contestants with different abilities. Additionally, it puts contestants' strategies in terms of what their opponents will observe. With this in mind, we write contestants' strategies and payoffs in terms of output and describe the effort of players only in the context of providing intuition for the results. Expected payoffs for contestant i in a single contest are

$$E[\pi^i(x^i, a^i)] = \Pr(x^{-i} < x^i) + \frac{1}{2}\Pr(x^{-i} = x^i) - c(x^i/a^i), \text{ for } i = 1, 2.$$

In the repeated contest game, contestant i maximizes the sum of expected payoffs in the two contests by choosing output levels in each contest, x_1^i and x_2^i. For a given strategy of player $-i$, we express the expected payoffs of player i over the two contests below.

$$\begin{aligned} E[\pi^i(x_1^i, x_2^i, a^i)] &= E[\pi_1^i(x_1^i, a^i)] + E[\pi_2^i(x_2^i, a^i)|\mu^{-i}(x_1^i)] \\ &= \Pr(x_1^{-i} < x_1^i) + \frac{1}{2}\Pr(x_1^{-i} = x_1^i) - c\left(x_1^i/a^i\right) \\ &+ \Pr(x_2^{-i} < x_2^i|\mu^{-i}(x_1^i)) + \frac{1}{2}\Pr(x_2^{-i} = x_2^i|\mu^{-i}(x_1^i)) - c(x_2^i/a^i) \text{ for } i = 1, 2. \end{aligned}$$

In the following sections, we will derive the properties of the unique symmetric perfect Bayes Nash equilibrium in the repeated contest game.

[7] Equivalent to the notion of private information about ability is private information about the cost of output. Additionally, if the cost of effort is linear, then this framework is equivalent to an all-pay auction where values are private information and bids are observed

3 Second Contest

After the first of two contests, each contestant will believe that their opponent has high ability with some probability. For each set of these probabilities, the equilibrium characterized in this section will be played in the second contest. Therefore, the expected payoffs of the contestants in this section will be equal to the continuation payoffs of the second contest in any perfect Bayesian equilibrium of the repeated contest game. Finding the equilibrium strategies for the second contest should be viewed as the first step in finding the standard backwards induction solution of the model that was introduced in the previous section.

In this section, we name our two contestants the strong contestant and the weak contestant, so that $i = s, w$ and $\mu_s \geq \mu_w$, where $\mu_i = \Pr(a^i = a_h)$. Ex-ante, the strong contestant, is at least as likely to have high ability as the weaker contestant. However, this does not rule out the possibility of the weaker contestant having high ability or the stronger contestant having low ability, or both.

3.1 Strategies

The strategies of each contestant consist of output distributions for both low and high ability realizations. We define $L_i(x|\mu_i, \mu_{-i}) \equiv \Pr(x^i \leq x|a^i = a_\ell, \mu_i, \mu_{-i})$ and $H_i(x|\mu_i, \mu_{-i}) \equiv \Pr(x^i \leq x|a^i = a_h, \mu_i, \mu_{-i})$ which denote these respective distributions. The ex-ante output distribution of contestant i, defined as $F_i(x|\mu_i, \mu_{-i}) \equiv \Pr(x^i \leq x|\mu_i, \mu_{-i})$, also represents the ex-interim output distribution of player i from the perspective of player $-i$. Consistency of information sets requires $F_i(x|\mu_i, \mu_{-i}) = (1 - \mu_i)L_i(x|\mu_i, \mu_{-i}) + \mu_i H_i(x|\mu_i, \mu_{-i})$. Additionally, let $\ell_i(x|\mu_i, \mu_{-i})$, $h_i(x|\mu_i, \mu_{-i})$ and $f_i(x|\mu_i, \mu_{-i})$ be the densities induced from $L_i(x|\mu_i, \mu_{-i})$, $H_i(x|\mu_i, \mu_{-i})$ and $F_i(x|\mu_i, \mu_{-i})$.[8] For simplicity, we suppress the probabilities, (μ_i, μ_{-i}), from the notation of the output distributions for the remainder of this section.

Given an expected output distribution of their opponent, the best response set for contestant i with ability a^i is

$$BR_i(a^i) \equiv \{x : E[\pi^i(x^i, a^i)] \geq E[\pi^i(\tilde{x}^i, a^i)], \forall \tilde{x}^i \geq 0\}.$$

The support of contestants' strategies are denoted by $X_\ell^i \equiv \{x : \ell_i(x) > 0\}$ and $X_h^i \equiv \{x : h_i(x) > 0\}$. An equilibrium is a set of output distributions, $(L_s(x), H_s(x), L_w(x), H_w(x))$, such that $X_\ell^i \subseteq BR_i(a_\ell)$, and $X_h^i \subseteq BR_i(a_h)$ for $i = s, w$. The general properties of an equilibrium are outlined in the following lemma; the proof is in the appendix.

Lemma 3.1. *In any equilibrium, contestants' distributions of output, $H_s(x), L_s(x), H_w(x)$, and $L_w(x)$, are continuous on $(0, x^*)$, where*

$$x^* \equiv \sup\{BR_s(a_\ell) \cup BR_s(a_h)\} = \sup\{BR_w(a_\ell) \cup BR_w(a_h)\}, \text{ and}$$
$$0 = \inf\{BR_s(a_\ell) \cup BR_s(a_h)\} = \inf\{BR_w(a_\ell) \cup BR_w(a_h)\}$$

[8]Here we use the extended definition of density using Dirac-delta functions where necessary to properly define these densities when their corresponding distributions have mass points.

The combined best response sets of low ability and high ability types must be an interval for both the strong and weak contestants. Moreover, the interval is the same for each contestant. Intuitively, if the supremum of the interval was smaller for one of the contestants, then the other contestant would be wasting effort by sometimes producing more output than would ever be necessary to win the contest. Additionally, if there are gaps of positive measure in this combined interval for either contestant, then the opponent would have no incentive to produce output in the interior of the gap. This leads to a gap in the best response intervals for both players which cannot happen in equilibrium. This argument, which is formalized in the proof of Lemma 3.1, implies that the infimum of the best response interval of each player must be zero. Lastly, continuity of the output distributions implies no positive output is chosen with positive probability. [9]

While the fundamentals of this model are somewhat different to those studied by Siegel (2014), the general properties of the equilibrium strategies are the same. In his all-pay auction setting, when types are independently drawn and the value of winning the contest increases in type, he shows that there is a unique monotonic equilibrium, i.e., for each contestant, all bids of the high type are at least as large as all bids of the low type. These properties hold in our setting where abilities act as types and outputs are chosen rather than bids.[10] This implies that there is a unique equilibrium where for $i = s, w$ and any $x \in BR_i(a_h)$ and $x' \in BR_i(a_\ell)$ it must be that $x' < x$. This fact, combined with the previous lemma implies that $\sup\{BR_i(a_\ell)\} = \inf\{BR_i(a_h)\}$. This cut-off is denoted by x_i^* for $i = s, w$.

Proposition 3.2 (Unique Equilibrium - Single Contest)**.** *There is a unique equilibrium,* $(L_s^*(x), H_s^*(x), L_w^*(x), H_w^*(x))$*, where* $\overline{X_i^\ell} = BR_i(a_\ell) = [0, x_i^*]$*,* $\overline{X_i^h} = BR_i(a_h) = [x_i^*, x^*]$ *for* $i = s, w$ *and* $0 \leq x_s^* \leq x_w^* \leq x^*$*.*

Here we highlight the important details of the construction of the equilibrium; see appendix A for the technical details and the algebraic form of the strategies.

First, from Lemma 3.1, the combined best response sets for the strong contestant and the weak contestant are the same, which we denote by the interval, $[0, x^*]$. Second, since the equilibrium is monotonic, the best response sets of each ability are disjoint for each contestant, with the set for high ability ranging over larger outputs than the set for low ability. Third, since the strong player is more likely to be high ability, the length of the best response set of high ability for is longer for the strong player. The basic structure of these best response sets is shown in Figure 1.

[9]This does not preclude one of the contestants from choosing zero output with positive probability.

[10]While, in this section, we borrow heavily from the properties of Siegel (2014) and follow his construction to characterize the unique equilibrium, the setting is slightly different and therefore we cannot directly use his results. His contestants differ on value of the prize rather than ability and the cost of output, which is a simple bid, is linear. Monotonicity holds in the current model as a player's marginal cost for a given output is ranked by type, where his contestants marginal value is ranked by type for all bids. A simple transformation makes these two settings isomorphic.

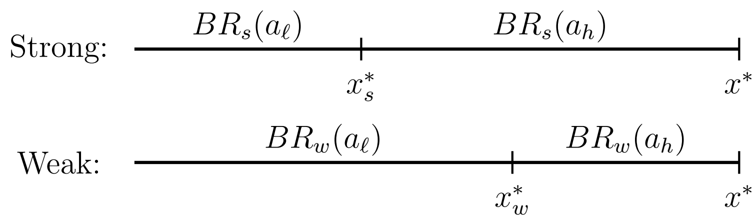

Figure 1: Representation of best response sets of the strong and weak players.

Characterizing the output distributions of each contestant in the unique equilibrium requires using an indifference condition for each contestant in each best response set. Starting from x^* and working backwards toward zero, these indifference conditions pin down the value of x^* and subsequently x_w^* and x_s^*.

For example, each contestant must be indifferent between producing any output level from x^* to x_w^* when they have high ability. This means that the marginal benefit of increasing output in this range must equal the marginal cost. This benefit is the increased chance of winning the contest, namely the density of the output distribution of the opponent. Therefore $f_i(x) = c'(x/a_h)/a_h$ for $x \in (x_w^*, x^*)$, $i = s, w$, and the output distributions over this range are the same for both contestants.

The expected output distribution of the weak and strong players are

$$F_s^*(x) = \begin{cases} c(x), & 0 \le x \le x_w^* \\ 1 - \mu_w - c\left(\frac{c^{-1}(1-\mu_w)}{a_h}\right) + c\left(\frac{x}{a_h}\right), & x_w^* \le x \le x^* \end{cases}$$

$$F_w^*(x) = \begin{cases} F_w^*(0) + c(x), & 0 \le x \le x_s^* \\ 1 - \mu_w - c\left(\frac{c^{-1}(1-\mu_w)}{a_h}\right) + c\left(\frac{x}{a_h}\right), & x_s^* \le x \le x^* \end{cases}$$

An example of these distributions with $a_h = 2$ and $c(e) = \frac{1}{2}e^2$ is depicted below.

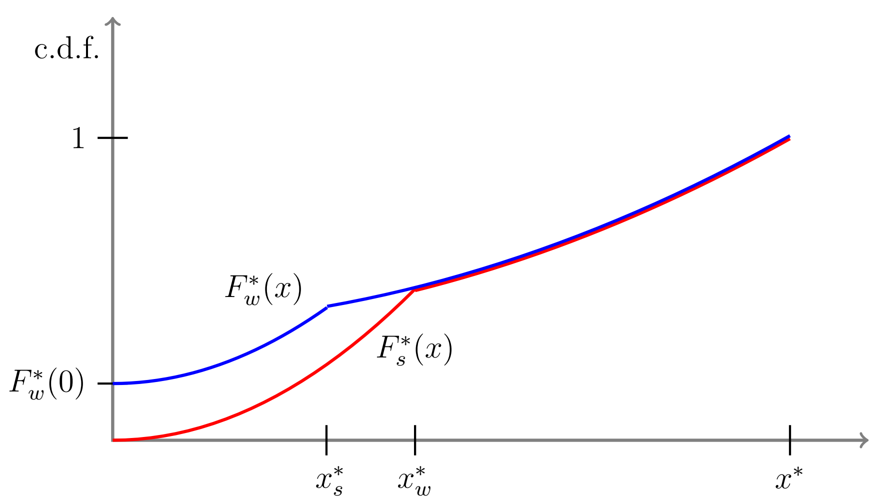

Figure 2: Expected output distributions of contestants in a single contest.

Because the best response sets of low and high ability types are disjoint for each contestant, we can recover the output distributions for each type of each contestant.

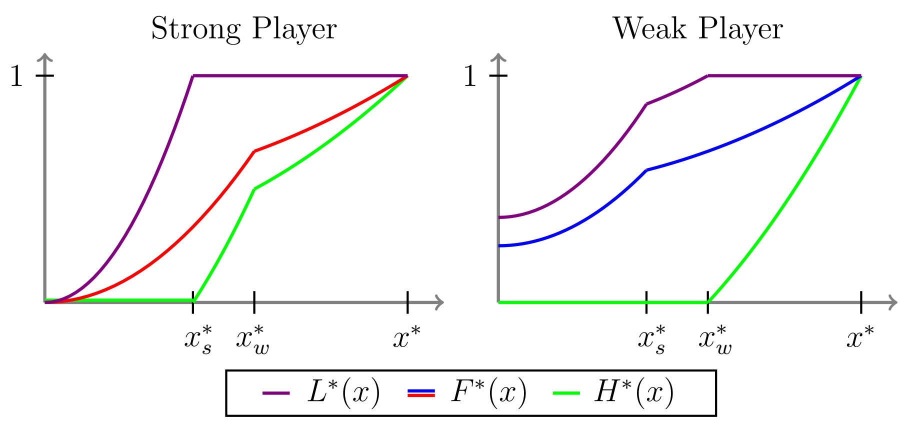

Figure 3: Output distributions for high and low ability type of the strong and weak player in a single contest. ($a_h = 2$ and $c(e) = \frac{1}{2}e^2$)

3.2 Payoffs

For strategies in the first contest, the main objects of interest from the equilibrium of the second contest are the payoffs of the contestants. Since the equilibrium for any pair of probabilities (μ_s, μ_w) is unique, these will equal the payoffs that contestants expect to receive in the second contest for any set of beliefs that result from the first contest. The expected payoffs are functions of a contestant's ability and the probabilities of each contestant being high ability as viewed by their opponent. These are denoted as $v_i(\mu_i, \mu_{-i}, a^i) = E[\pi^i(\hat{x}^i, a^i)]$ where $\hat{x}^i \in BR_i(a^i)$. Characterization of the payoffs are in Corollary 3.3.

Corollary 3.3. *The ex-interim expected payoff of each contestant is*

$$v_s(\mu_s, \mu_w, a_h) = v_w(\mu_w, \mu_s, a_h) = 1 - \mu_w - c\left(\frac{c^{-1}(1-\mu_w)}{a_h}\right)$$

$$v_s(\mu_s, \mu_w, a_\ell) = \mu_s - \mu_w - \left[c\left(\frac{c^{-1}(1-\mu_w)}{a_h}\right) - c\left(\frac{c^{-1}(1-\mu_s)}{a_h}\right)\right]$$

$$v_w(\mu_w, \mu_s, a_\ell) = 0.$$

For both the strong and weak contestants who are high ability, the expected payoff is entirely determined by the value of x^*. This value is pinned down by the ex-ante expected output distribution of the stronger contestant which, in turn, is constructed from the indifference conditions of the weaker contestant. Therefore, x^* is a determined entirely by μ_w, the probability that the weaker contestant has high ability. Intuitively, high ability contestants are confident they can win, but the overall competitiveness of the contest will determine how much effort they need to exert to do so. This payoff decreases as μ_w increases, implying that increased competition will increase the effort of high ability contestants, decreasing their expected payoff.

For contestants who are low ability, expected payoffs are determined by how often they can win a contest with arbitrarily low effort. The strong contestant will exert no effort with probability zero, while the weak contestant will exert no effort with a probability that increases with the strength of the relative strength of their opponent. Intuitively, a low ability contestant becomes discouraged when he believes that his opponent is likely to have high ability. Therefore, the low ability type of the weaker contestant will never win a contest when they exert no effort, leading to an expected payoff of zero. The stronger contestant who has low ability will have positive expected payoffs which increase with the contestant's relative strength.

How these payoffs change with respect to the probabilities of each contestant having high ability are important for the analysis of strategies in the first contest. Namely, the contestants can affect their perceived strength in this second contest through their choice of output in the first contest. For a contestant with high ability, payoffs decrease when the contest appears more competitive. Therefore, they prefer to look weak entering the second contest in order to reduce the perceived level of competition. On the other hand, the payoffs of low ability contestants increase when they appear strong to their opponent. These countervailing incentives, formalized in the proposition below, are a significant strategic force in the first contest.

Proposition 3.4 (Countervailing incentives)**.** *For all $\mu_i \in (0,1)$, expected payoffs in the second contest decrease for high ability players as μ_i increases, $\frac{\partial}{\partial \mu_i} E[v_i(\mu_i, \mu_{-i}, a_h)] < 0$, and increase with μ_i for low ability players, $\frac{\partial}{\partial \mu_i} E[v_i(\mu_i, \mu_{-i}, a_\ell)] > 0$.*

In particular, the marginal effect of beliefs on payoffs in the second contest is given by

$$\frac{\partial}{\partial \mu_i} E[v_i(\mu_i, \mu_{-i}, a_h)] = d(\mu_i)(F_{\mu_{-i}}(\mu_i) - 1) < 0 \text{ and}$$

$$\frac{\partial}{\partial \mu_i} E[v_i(\mu_i, \mu_{-i}, a_\ell)] = d(\mu_i) F_{\mu_{-i}}(\mu_i) > 0,$$

$$\text{where } d(\mu_i) \equiv \left[1 + \frac{\partial}{\partial \mu_i} c\left(\frac{c^{-1}(1-\mu_i)}{a_h}\right)\right] \text{ and } F_{\mu_{-i}}(\mu_i) = \Pr(\mu_{-i} \leq \mu_i).$$

The convexity of the cost of effort implies that $d(\mu_i) \in \left[\frac{a_h - 1}{a_h}, 1\right)$ for all μ_i, which guarantees that these incentives are strict.

4 First contest

We now turn to the analysis of the first contest. Prior to this contest, the contestants are ex-ante symmetric with equal chance of being high or low ability. They become privately informed of their ability prior to choosing output in the first contest. Additionally, the contestants are forward looking, understanding that their output choice in the first contest will be revealed to the other contestant and subsequently will effect that contestant's beliefs and therefore the second contest strategies as analyzed in the previous section.

4.1 Strategies

For each player $i = 1, 2$, $L_1^i(x)$ and $H_1^i(x)$ denote the first period output distributions of a contestant with low ability and high ability respectively. Then the ex-ante expected output distribution is $F_1^i(x_1) = \frac{1}{2}L_1^i(x_1) + \frac{1}{2}H_1^i(x_1)$, for $i = 1, 2$. Additionally, f_1^i, ℓ_1^i and h_1^i denote the densities that are induced from the distribution functions F_1^i, L_1^i and H_1^i.[11] Lastly, define $X_1^{h,i} = \{x | h_1^i(x) > 0\}$ and $X_1^{\ell,i} = \{x | \ell_1^i(x) > 0\}$. Since contestants are symmetric, we restrict attention to equilibria that are symmetric.

A set of output distributions $\{H_1^i(x_1), L_1^i(x_1), H_2^i(x_2|\mu_i, \mu_{-i}), L_2^i(x_2|\mu_i, \mu_{-i})$ for $i = 1, 2\}$ form a symmetric perfect Bayesian equilibrium (SPBE) for two successive contests if

1. strategies are symmetric: $H_1^1(x) = H_1^2(x)$, $L_1^1(x) = L_1^2(x)$, $H_2^1(x|\mu_1, \mu_2) = H_2^2(x|\mu_2, \mu_1)$, and $L_2^1(x|\mu_1, \mu_2) = L_2^2(x|\mu_2, \mu_1)$,

2. contestants play the unique Bayes Nash equilibrium in the second contest: for $i = 1, 2$ and for every (μ_i, μ_{-i}),

$$(H_2^i(x|\mu_i, \mu_{-i}), L_2^i(x|\mu_i, \mu_{-i})) = \begin{cases} (H_w^*(x|\mu_i, \mu_{-i}), L_w^*(x|\mu_i, \mu_{-i})), & \text{if } \mu_i \le \mu_{-i} \\ (H_s^*(x|\mu_i, \mu_{-i}), L_s^*(x|\mu_i, \mu_{-i})), & \text{if } \mu_i > \mu_{-i} \end{cases}$$

3. players update beliefs according to Bayes rule when feasible:[12]

$$\mu_i = \mu(x_1^i) = \frac{h_1(x_1^i)}{h_1(x_1^i) + \ell_1(x_1^i)}, \text{ for } i = 1, 2, \text{ and}$$

4. given (2) and (3), contestants always choose an optimal output in the first contest: for $i = 1, 2$, for every $x_1^i \in X_1^{\ell,i}$ player i chooses an

$$x_1^i \in \arg\max_{x^i} E[\pi(x_1^i, x_2^i(\mu(x_1^i), \mu(x_1^{-i}), a_\ell), a_\ell)] \equiv BR_i(a_\ell),$$

and for every $x_1^i \in X_1^{h,i}$ player i chooses an

$$x_1^i \in \arg\max_{x^i} E[\pi(x_1^i, x_2^i(\mu(x_1^i), \mu(x_1^{-i}), a_h), a_h)] \equiv BR_i(a_h).$$

The output choice of contestants in the first contest determine payoffs in that contest and their opponent's belief about their ability in the second contest. As was characterized in the previous section, the strategies in the second contest that satisfy condition (2) and the contestants' respective expected payoffs in that contest depend on these beliefs. In particular the payoffs are $v_i(\mu(x_1^i), \mu(x_1^{-i}), a^i)$ as in Corollary 3.3, where $i = s$ if $\mu(x_1^i) \ge \mu(x_1^{-i})$ and $i = w$ otherwise. Therefore the expected payoffs to player i for the two contests can be written in terms of output in the first contest:

$$E[\pi^i(x_1^i, \hat{x}_2^i(\mu(x_1^i), \mu(x_1^{-i}), a^i), a^i)] = E[\pi_1^i(x_1^i, a^i)] + E[v_i(\mu(x_1^i), \mu(x_1^{-i}), a^i)].$$

[11] Again the extended definition of density using Dirac-delta functions is invoked where necessary.

[12] Using the extended definition of density allows agents to update their beliefs even when they see their opponent produce an output where the distribution has a mass point. For example, if the H_1 has a mass point at x, while L_1 does not, this definition implies $\mu(x_1) = 1$.

Differences in marginal cost of output for the first contest and the incentives stemming from the second contest combine to require that the equilibrium belief function, $\mu(x)$, is non-decreasing in first period output. Supposing the opposite implies that a higher output would result in a lower belief about the ability of the contestant. From conditions (3) and (4), the higher output would be in the $BR(a_\ell)$ and the lower output would be in $BR(a_h)$. However, the low ability contestant must strictly prefer the lower output to the higher output due to the higher marginal costs of this contestant and the higher expected payoffs in the second contest from appearing stronger. The proof of this lemma and other results of this section are relegated to the appendix.

Lemma 4.1 (Monotonic beliefs)**.** *In every SPBE, $\mu(x)$ is weakly increasing in x for all $x \in X_1 = X_1^h \cup X_1^\ell$.*

In addition to restricting the belief function on the equilibrium path, the countervailing incentives also restrict the beliefs that can be assumed for outputs that are not chosen in equilibrium. In games where players can signal private information, there are often equilibria where actions off the equilibrium path are not taken because players who do so are assumed to have negative characteristics.[13] However, any off-path belief will benefit at least one type of contestant due to their countervailing incentives. Because the belief function must be well behaved for all potential outputs in the first contest, the first period strategies conform to nice properties in equilibrium. In particular, equilibrium strategies in the first contest have no atoms, implying that payoffs in the first contest are continuous in output, and there are no gaps in the corresponding best response sets. These properties are described in the following lemmas.

Lemma 4.2. *There is no output that is played with positive probability and* $\Pr(win|x) = F_1(x)$ *is continuous.*

Lemma 4.3. *$BR(a_\ell)$ and $BR(a_h)$ are intervals where $0 = x_{\ell,*} \leq x_{h,*} < x_\ell^* \leq x_h^*$ and we define $x_{\ell,*} = \inf\{BR(a_\ell)\}, x_\ell^* = \sup\{BR(a_\ell)\}, x_{h,*} = \inf\{BR(a_h)\}$ and $x_h^* = \sup\{BR(a_h)\}$.*

As shown in Lemma 4.3, the countervailing incentives also guarantee that the intersection of the best response sets for the low ability and high ability contestants, $BR(a_\ell) \cap BR(a_h)$, is non-trivial. Without gaps between the best response sets, the only alternative construction would involve increasing best response sets which only share an endpoint, analogous to the best response sets of each contestant in the second contest. This would not be on equilibrium strategy as both high and low ability contestants would prefer to deviate to the other best response set near the endpoint. Both types of contestants can significantly increase their expected payoff in the second contest by mimicking the other type while only bearing a small cost in the first contest.

In any equilibrium, the best response sets of the high and low ability contestants partition outputs of the first contest into three distinct intervals. These intervals and best response sets are shown in Figure 4.

[13] See Spence (1973).

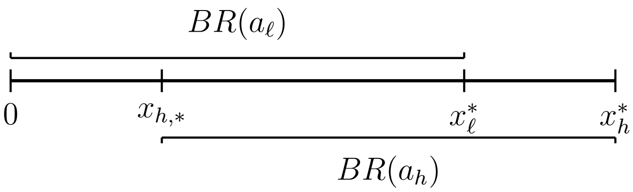

Figure 4: Representation best response sets of the high ability and low ability contestants in the first contest.

The algebraic conditions for output being in $BR(a_h)$ and $BR(a_\ell)$ respectively are

$$BR(a_h): \; F_1^*(x) + E[v_i(\mu(x), \mu(x^{-i}), a_h)] - c\left(\frac{x}{a_h}\right) = k_h \text{ and}$$

$$BR(a_\ell): \; F_1^*(x) + E[v_i(\mu(x), \mu(x^{-i}), a_\ell)] - c(x) = k_\ell,$$

where k_h and k_ℓ are the expected payoffs of high and low ability contestants in an equilibrium of the repeated contest game.

Given the properties of the strategies and best response sets of the first contest, it can be shown that the belief function $\mu(x)$ must be continuous on the union of the best response sets, $[0, x_h^*]$. First, a contestant's combined expected payoffs for the two contests is constant for all outputs chosen in the first contest within the contestant's respective best response set. Additionally, cost of effort is continuous in output, and from Lemma 4.2, the payoff functions in the first contest are continuous in output. This implies that expected payoffs in the second contest must also be continuous in first contest output. Lastly, as shown in Proposition 3.4, payoffs in the second contest are strictly increasing (decreasing) in the belief function, $\mu(x)$, for low (high) ability contestants, and therefore $\mu(x)$ must also be continuous in output. This argument and its implications on the belief function are formalized in Lemma 4.4.

Lemma 4.4. *The belief function and the distribution functions of output, $L_1(x)$ and $H_1(x)$, are continuous in output on $[0, x_h^*]$. Additionally, the belief function is given by $\mu(x) = 0$ for all $x \in [0, x_{h,*}]$, $\mu(x) = 1$ for all $x \in [x_\ell^*, x_h^*]$ and is weakly increasing on $(x_{h,*}, x_\ell^*)$.*

The ex-ante output distribution for each contestant in the first contest can be characterized in each of the three regions. For outputs in the range of $0 \leq x < x_{h,*}$, Lemma 4.5 states $\mu(x) = 0$ and therefore $E[v_i(\mu(x), \mu(x^{-i}), a_\ell)] = 0$. Additionally, since $x = 0$ is in $BR(a_\ell)$, $k_\ell = 0$ in any equilibrium. Therefore from the condition for $BR(a_\ell)$, the output distribution equals the cost function of output for low ability contestants, $F_1^*(x) = c(x)$. Similarly, for outputs in the range $x_\ell^* < x \leq x_h^*$, $E[v_i(\mu(x), \mu(x^{-i}), a_h)] = v_h$ and the distribution function is $F_1^*(x) + v_h = c(x/a_h) + k_h$, where $v_h = E[v_i(1, \mu(x_j), a_h)]$, the expected payoff in the second contest for a contestant who is known to have high ability.

All outputs in the range $x_{h,*} \leq x \leq x_\ell^*$ must be in the best response set of both low and high ability contestants. Therefore, both low and high ability contestants are indifferent between choosing any output in this range. However, the marginal cost of increasing output for the low ability contestant is always more than for the high ability contestant. This can only be true if increasing output benefits the low ability contestant more than one with high ability. Because the benefit of increasing output in the first contest is identical in the first contest, this difference must come from the expected payoffs in the second contest. Therefore, the difference in marginal benefits of increasing output, and thus appearing stronger in the second contest must equal the difference in marginal costs that they face in the first contest. This condition determines the derivative of the belief function over this range of outputs. Note that the marginal benefits of looking stronger in the second contest for both low and high ability contestants are given in Proposition 4.1.

$$\mu'(x)d(\mu(x)) = c'(x) - \frac{1}{a_h}c'(x/a_h)$$

Combining this with the best response conditions of both the low and high ability contestants results in a differential equation that the output density function must satisfy over this interval.

$$f_1^*(x) = \frac{\partial}{\partial x}c(x)(1 - F_1^*(x)) + \frac{\partial}{\partial x}c\left(\frac{x}{a_h}\right)F_1^*(x) \quad (\dagger)$$

Moreover, due to the continuity result of Lemma 4.2, the ex-ante output distribution function over this interval is uniquely determined by the value of $x_{h,*}$.

To show uniqueness of strategies in the first contest, it suffices to show $x_{h,*}$ is unique. This endpoint determines the ex-ante output distribution function on $BR(a_\ell)$ and in turn the belief function over this same set. These two functions can be used to calculate the output distribution function of both low and high ability contestants. Lastly, the remaining values, x_ℓ^* and x_h^* are pinned down respectively by these two distribution functions. The uniqueness of $x_{h,*}$ along with the details of this argument are given in the proof of Theorem 4.5.

Theorem 4.5 (Uniqueness of equilibrium). *There is a unique symmetric perfect Bayes Nash equilibrium* $\{(L_1^*(x_1), L_2^*(x_2|\mu_i, \mu_{-i})), (H_1^*(x_1), H_2^*(x_2|\mu_i, \mu_{-i}))\}$.

4.2 Information and expected payoffs

Observed output in the first contest from each contestant will land in one of three intervals. If the output is between 0 and $x_{h,*}$, then the contestant is revealed to have low ability, while output between x_ℓ^* and x_h^* must have been produced by a contestant with high ability. Output between these ranges may have been produced by a contestant with either high or low ability.

Low ability contestants either choose a low effort which produces an output that reveals their ability to their opponent, or they decide to *bluff* by choosing a higher effort which produces an output that may have been produced by a high ability contestant. This higher effort level will provide a lower expected payoff in the first contest as the additional cost of effort will exceed the benefit of increasing the probability of winning this contest. These contestants are willing to put in the additional effort to have a stronger position entering the second contest, as they benefit from appearing to have high ability. Therefore ex-interim expected payoffs for a low ability contestant are negative in the first contest and positive in the second contest.

High ability contestants either choose a high effort which produces an output that reveals their ability to their opponent, or they decide to *sandbag* by choosing a lower effort which produces an output that may have been produced by a low ability contestant. Since effort is less costly to high ability contestants, the lower cost of reduced effort will not offset the lower winning probability of the first contest. These contestants are willing to produce the lower output in the first contest as they benefit from appearing to have a low ability entering the second contest. The per period expected payoffs are depicted in Figure 5.

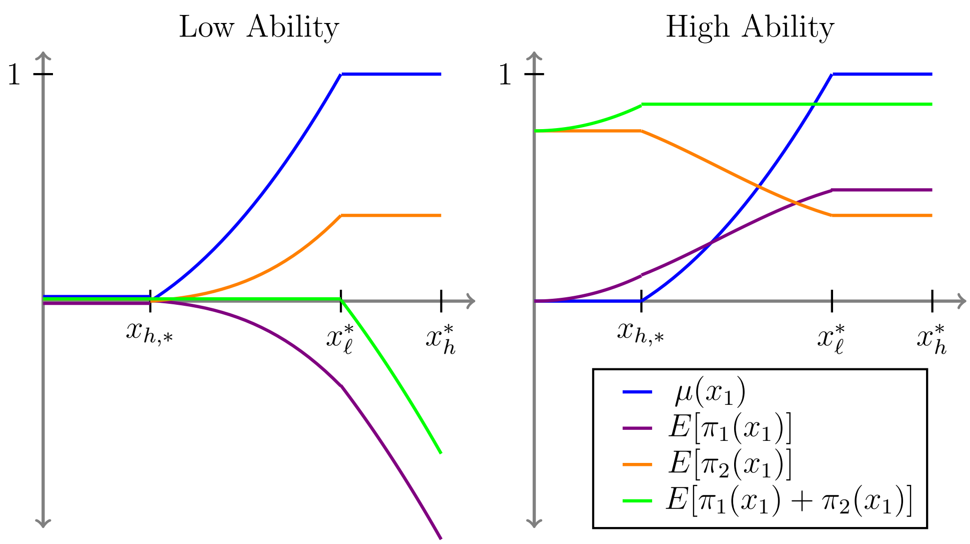

Figure 5: Beliefs and expected payoffs as a function of first period output ($c(e) = \frac{1}{2}e^2$ and $a_h = 2$).

4.3 Equilibrium construction

In addition to showing uniqueness of equilibrium given any weakly convex cost function, this equilibrium can be constructed given that the cost function of effort is a monomial, i.e. $c(x) = kx^\alpha$, with $\alpha \geq 1$ and $k > 0$. Note that this includes the linear cost case of the all-pay auction. A graphical representation of the equilibrium output functions for quadratic cost of effort is given in Figure 6. A general method for construction is given in appendix A.

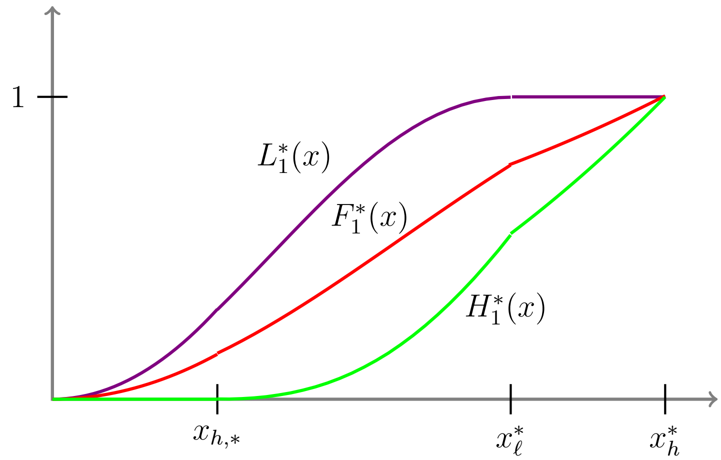

Figure 6: Distribution of strategies in the first of two successive contests.
$(c(e) = \frac{1}{2}e^2,\ a_h = 2)$

5 Welfare

Relevant properties for contest design include both the output and payoffs of contestants as well as the outcome of the two contests. In order to analyze the strategic effect of private information in a dynamic contest, we compare these properties in the repeated contest model to a benchmark where contestants compete with the same payoff structure but the possibility of signaling private information through actions is suppressed. In particular, the expected outcomes along with the payoffs and outputs of contestants in the second contest with ex-ante symmetric contestants ($\mu_s = \mu_w = 1/2$) can be doubled and used as this benchmark.

This can be interpreted in two ways. First, as two separate contests where contestant outputs and the contest winner are not revealed after the first contest. Instead, the winner of each contest is revealed only after the second contest has ended. Second, as one longer contest, where either the cost function is stretched by factor of two, or cost is a function of the intensity of effort chosen by the contestant in each period of the contest.

5.1 Outcomes

In the repeated contest equilibrium, overlapping best response sets give a low ability contestant a positive probability of beating a high ability contestants in the first contest. Additionally, if the low ability contestant enters the second contest in a stronger position, which is always the case when they win the first contest, they may also win the second contest. In the benchmark, the best response sets for high and low ability contestants are disjoint, implying that a high ability contestant will always win a contest against a low ability contestant. This property is used to help motivate use of multi-stage tournaments by both Moldovanu and Sela (2006) and Fullerton and McAfee (1999). Additionally, Ye (2007) shows that an all-pay bidding stage can help guarantee efficient entry into a first price auction that has a high participation cost. Our results show that if contestants can learn about future opponents from choices in the first stage of a tournament, then the best participant may not win the tournament, and in fact, may not make it to the second stage.

To see this connection, consider a four player, two stage tournament where the payoffs of each stage is identical to each contest in the current model and output from the first stage is observed by all four players before the second stage. Because these outputs are sufficient to characterize the second round strategies, the countervailing incentives in the first round will guarantee partial pooling in the first stage allowing for the possibility of a low ability contestant beating a high ability contestant in that first stage. Non-existence of a separating equilibrium in the first round of this four player tournament has been shown by Zhang and Wang (2009) in the all-pay auction setting with ex-ante symmetric contestants who have values of winning the tournament distributed on a compact support.

5.2 Payoffs and output

In order to compare total expected output of the contestants and their respective expected payoffs we need to construct the equilibrium of the repeated contest model. We therefore assume the cost of effort is $c(e) = ke^{\alpha}$, with $\alpha \geq 1$ and $k > 0$ and compare the strategies derived in section 3 when the type distribution is $\mu_s = \mu_w = \frac{1}{2}$ with the strategies of the repeated contest mentioned in section 4.3, which are detailed in appendix A.

Strategic incentives will reduce the expected output of high ability contestants and increase the expected output of low ability contestants in the first contest as compared to the expected output of each type of contestant in a single contest. Specifically, the output distribution of a high ability contestant is first order stochastically dominated by the high ability contestant's output distribution in the benchmark. On the other hand, the output distribution of a low ability contestant dominates the respective output distribution in the benchmark. This is shown in Figure 7. When the abilities of contestants are sufficiently different, the overall expected output is lower than in the benchmark as the effect of a high ability contestant outweighs that of a low ability contestant due to the larger impact of effort on output.

Compared to the benchmark, the expected aggregate output in the second contest is also lower. With high probability, one contestant will enter the second contest in a stronger position than the other. The difference in positions will reduce competition and therefore overall output. The weaker contestant faces a negative motivation effect, while the stronger contestant will not compete as aggressively against a weaker opponent. Combined with strategies in the first contest, this leads to lower expected aggregate output and a higher expected payoffs for high ability contestants in the repeated contest equilibrium.

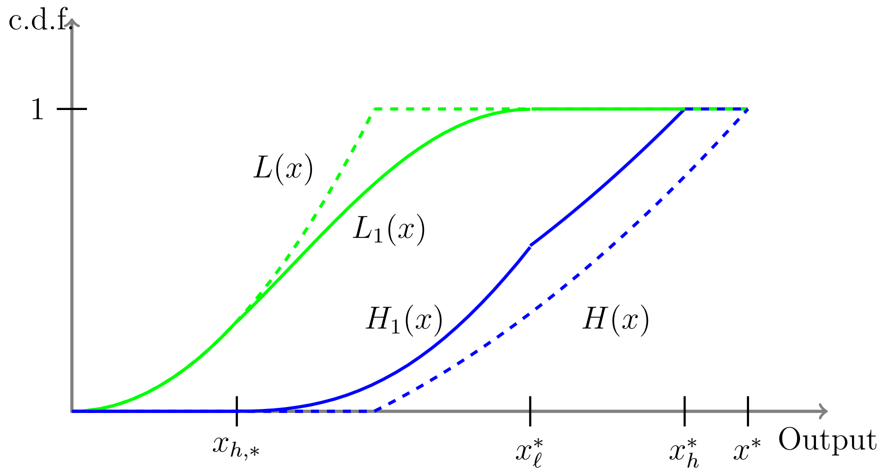

Figure 7: Distribution of strategies in the first of two contests compared to benchmark strategies. ($a_h = 2$ and $c(e) = \frac{1}{2}e^2$)

Theorem 5.1 (Increased aggregate payoffs)**.** *The expected payoff for the low ability contestant is the same per contest as the single contest benchmark, while the high ability contestant receives a higher expected payoff per contest.*

Proof. Equilibrium payoffs of a low ability contestant in the two contest model are 0. This is equal to the expected payoff in the single contest benchmark.

The expected payoff of a high ability contestant in the benchmark model is $\frac{1}{2A}$, where $A = \frac{a_h^\alpha}{a_h^\alpha - 1}$. Then over two contests without output revelation, the expected payoff for a high ability contestant is $\frac{1}{A}$. The expected payoff of a high ability contestant in repeated contest model with publicly revealed output, k_h, is higher:

$$k_h - \frac{1}{A} = \frac{1}{2A} + \frac{(2A-1)(1-a_h^\alpha e^{-1/A})}{2Aa_h^\alpha(1-e^{-1/A})} = \frac{1}{2A}\left(1 - \frac{(2A-1)(e^{-1/A} - 1/a_h^\alpha)}{1-e^{-1/A}}\right) > 0.$$

□

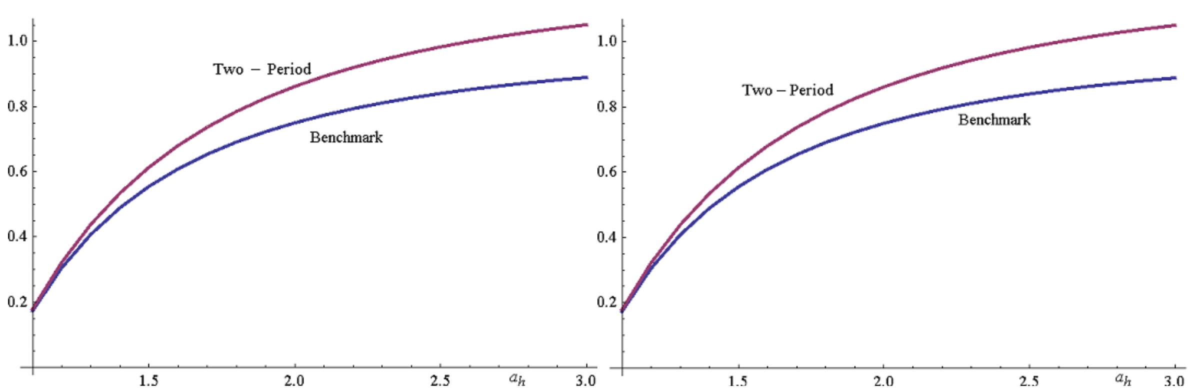

Figure 8: Payoffs of high ability contestant in terms of ability ratio, cost: $c(e) = e$ and $c(e) = e^2$.

Theorem 5.2 (Reduced aggregate output)**.** *Ex-ante expected aggregate effort of the contestants in each of the two contests is less than in the single contest.*

Proof. Ex-ante payoffs for the contestants are $\frac{k_h}{2}$ in the repeated contest, and $\frac{1}{2A}$ in the benchmark. Since the contestants are symmetric ex-ante, each will win each game with one half chance in both the repeated contest and benchmark models. Therefore, expected payoffs can be written as

$$E[\pi_1 + \pi_2] = 1 - E[c(x_1) + c(x_2)] = \frac{k_h}{2} > \frac{1}{2A} = 1 - E[2c(x)] = E[2\pi_b]$$

This implies that $E[c(x_1) + c(x_2)] < E[2c(x)]$. Also, since $c(\cdot)$ is weakly convex, then

$$E\left[2c\left(\frac{x_1 + x_2}{2}\right)\right] \leq E[c(x_1) + c(x_2)] < E[2c(x)].$$

Because $c(\cdot)$ is strictly increasing, this implies that

$$E[x_1 + x_2] < E[2x]$$

and therefore expected aggregate output in the repeated contest is less than twice the aggregate output in the benchmark. □

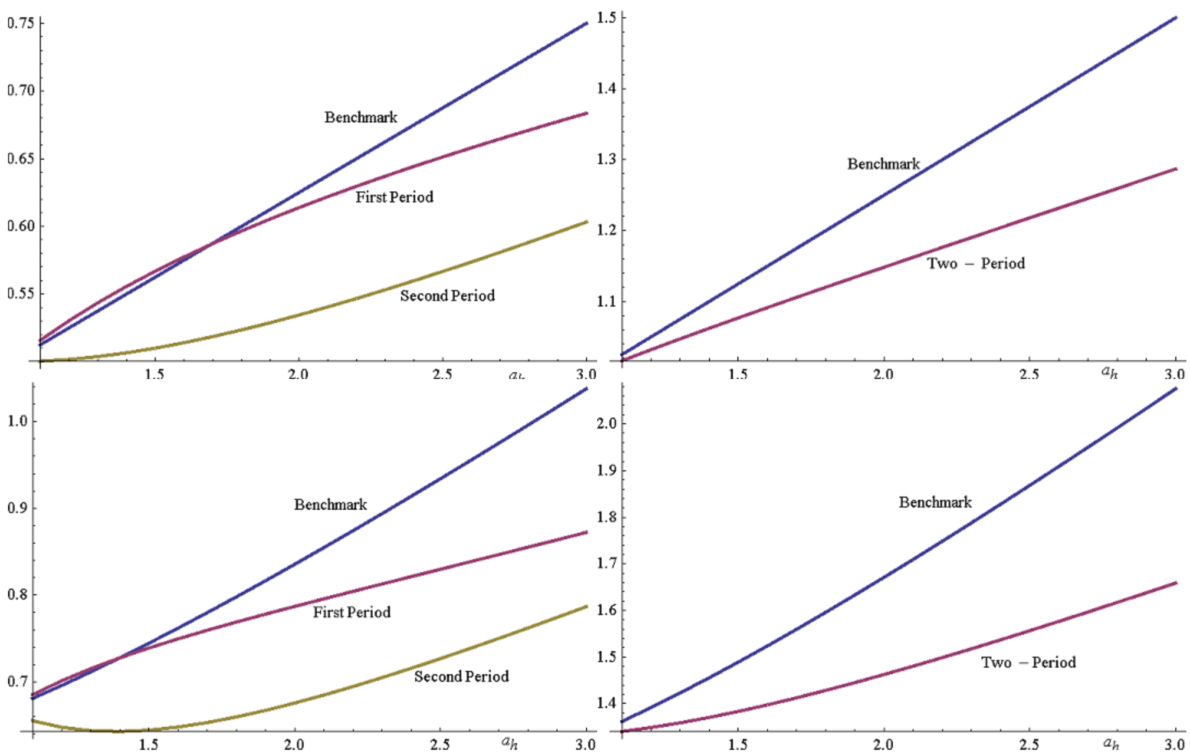

Figure 9: Output in terms of ability ratio, cost: $c(e) = e$ and $c(e) = e^2$.

5.2.1 Application to performance evaluations

Ederer (2010) and Aoyagi (2010) both discuss the potential merits of performance evaluations in a single contest between two employees. Performance evaluations divide one contest into two separate contests, where employees choose a level of effort before and after the evaluation. Aoyagi (2010) shows that when output is a noisy signal of effort and abilities do not effect output, performance evaluations reduce the expected output of the workers if the cost of effort is convex. On the other hand, Ederer (2010) shows that when ability affects output and the contestants do not know their ability, there are two competing effects of performance evaluations: the "evaluation effect" which stems from relative position in the contest and the "motivation effect" which which encourages the contestant who appears more productive. Strategically, the evaluation effect discourages the employee who is further behind while the motivation effect discourages employees who think they are less productive. The motivation effect also provides additional incentive for effort before the performance evaluation is administered because the employee wants to appear relatively more productive. This additional effort before the performance evaluation may outweigh the loss in output from the decreased competition after the evaluation.

When employees have private information about their abilities, the effect in the first period is not one directional. After the midterm evaluation, a high ability employee would actually prefer to look weaker and therefore will produce less effort before the evaluation. This could effectively counteract any additional effort exerted by low ability employees before the midterm evaluation. Additionally, after the evaluation, both differentiation between employees abilities and the discouragement effect stemming from one employee falling behind will combine to reduce expected output. Therefore, in this setting, performance evaluations may not encourage additional effort from employees.

6 Conclusion

While competing in two repeated contests with asymmetric information, contestants have an incentive to give up potential profits in the first contest to prevent revealing their private information. This leads to both bluffing and sandbagging in the first contest and can cause the following inefficiencies as compared to the single contest benchmark. First, a low ability contestant has a positive probability of winning both contests against a contestant who has high ability. Second, repeated contests have a lower expected output than the single contest, and additionally, the expected output of the second contest is lower than that of the first. While the results may seem overwhelmingly negative, in settings where the payoffs of competitors are of interest, our results are positive as ex-ante expected payoffs are higher for the contestants. Additionally, the intuitions developed here apply in more general dynamic settings where private information is valuable, and we leave that for future work.

Appendix

A Equilibrium Construction

Single contest

Because the equilibrium is monotonic, we know $BR_i(a_h) = x^*$ for $i = s, w$. Additionally, each contestant must be indifferent between all $x \in (x_i^*, x^*)$ when they have high ability. Given high ability these contestants have the same marginal cost of output, and therefore the density of the expected output of their opponents must also be the some for both indifference conditions to hold. Therefore, $f_s^*(x) = f_w^*(x)$ for $x \in (\max\{x_s^*, x_w^*\}, x^*)$ and $F_s^*(x^*) = F_w^*(x^*) = 1$. Since $f_i^*(x) = \mu_i h_i^*(x)$ for all $x \in (x_i^*, x^*)$, then $h_s^*(x) \leq h_w^*(x)$ for all $x \in (\max\{x_s^*, x_w^*\}, x^*)$. Also, $H_i(x_i^*) = 0$, which implies that $x_s^* \leq x_w^*$. Therefore, for the remainder of the construction, there are three intervals to consider: the best response set of the low types of both the stronger and the weaker contestants, $0 \leq x \leq x_s^*$, the best response set of the low type of the weaker contestant and the high type of the strong contestant, $x_s^* \leq x \leq x_w^*$, and best response set of the high types of each contestant, $x_w^* \leq x \leq x^*$.

Within their best response sets, contestants must be indifferent between all output levels. For example, contestant s given that he has ability of a_h, must be indifferent to picking all outputs between x_s^* and x^*. Then, for any for any x and x' in this interval the payoffs for the strong contestant must be the same. This puts a condition on $H_w(x)$, the output distribution of the weak contestant with high ability on the interval $[x_w^*, x^*]$, as the indifference for the strong contestant implies

$$\mu_w H_w^*(x') - c\left(\frac{x}{a_h}\right) = \mu_w H_w^*(x') - c\left(\frac{x'}{a_h}\right).$$

Rearranging and taking the limit as $x \to x'$,

$$\lim_{x \to x'} \frac{H_w^*(x) - H_w^*(x')}{c(\frac{x}{a_h}) - c(\frac{x'}{a_h})} = \frac{\partial H_w^*}{\partial c(\frac{x'}{a_h})} = \frac{1}{\mu_w}.$$

We use this to calculate the output density of the weak contestant on this interval.

$$\lim_{x \to x'} \frac{H_w^*(x) - H_w^*(x')}{x - x'} = \lim_{x \to x'} \frac{H_w^*(x) - H_w^*(x')}{c(\frac{x}{a_h}) - c(\frac{x'}{a_h})} \frac{c(\frac{x}{a_h}) - c(\frac{x'}{a_h})}{a_h(\frac{1}{a_h}(x - x'))}$$

$$h_w^*(x') = \frac{\partial H_w^*}{\partial c(\frac{x'}{a_h})} c'\left(\frac{x'}{a_h}\right) \frac{1}{a_h} = \frac{c'(x'/a_h)}{a_h \mu_w}$$

A similar calculation on each interval for each contestant allows us to characterize the densities of the output on each of the intervals below.

- $x_w^* \leq x \leq x^*$: $h_s^*(x) = \frac{c'(x/a_h)}{a_h \mu_s}$, $h_w^*(x) = \frac{c'(x/a_h)}{a_h \mu_w}$, $f_s^*(x) = f_w^*(x) = \frac{c'(x/a_h)}{a_h}$.
- $x_s^* \leq x \leq x_w^*$: $h_s^*(x) = \frac{c'(x)}{\mu_s}$, $\ell_w^*(x) = \frac{c'(x/a_h)}{a_h(1-\mu_w)}$, $f_s^*(x) = c'(x)$, $f_w^*(x) = \frac{c'(x/a_h)}{a_h}$.

- $0 \le x \le x_s^*$: $\ell_s^*(x) = \frac{c'(x)}{1-\mu_s}$, $\ell_w^*(x) = \frac{c'(x)}{1-\mu_w}$, $f_s^*(x) = f_w^*(x) = c'(x)$.

From the definition of the best response sets and the consistency of the contestants' information sets, the distribution of output for each contestant must satisfy

$$L_i^*(x_i^*) = 1, \quad H_i^*(x_i^*) = 0, \quad F_i^*(x_i^*) = 1 - \mu_i, \quad F_i^*(x^*) = 1.$$

To find the endpoints we look at the stronger contestant's distribution of strategies. This contestant does not choose zero effort with positive probability, and therefore $L_s^*(0) = 0$. Using $L_s^*(x_s^*) = 1$ and the definition of $\ell_s^*(x)$ on $[0, x_s^*]$, we calculate x_s^*.

$$\int_0^{x_s^*} \ell_s^*(x)dx = L_s^*(x_s^*) - L_s(0) = \frac{c(x_s^*)}{1-\mu_s} = 1$$

Then $c(x_s^*) = 1 - \mu_s$, so that $x_s^* = c^{-1}(1-\mu_s)$. Similarly, $x_w^* = c^{-1}(1-\mu_w)$. From these endpoints we can calculate x^*.

$$\int_{x_s^*}^{x_w^*} h_s^*(x)dx = \frac{c(x_w^*) - c(x_s^*)}{\mu_s} = \frac{(1-\mu_w) - (1-\mu_s)}{\mu_s} = \frac{\mu_s - \mu_w}{\mu_s}$$

$$\int_{x_w^*}^{x^*} h_s^*(x_s)dx = 1 - \frac{\mu_s - \mu_w}{\mu_s} = \frac{\mu_w}{\mu_s}$$

$$\int_{x_w^*}^{x^*} f_s^*(x_s)dx = c\left(\frac{x^*}{a_h}\right) - c\left(\frac{c^{-1}(1-\mu_w)}{a_h}\right) = \mu_w$$

$$x^* = a_h c^{-1}\left(\mu_w + c\left(\frac{c^{-1}(1-\mu_w)}{a_h}\right)\right)$$

Lastly, we can pin down the probability that the weaker contestant exerts no effort, and the unique equilibrium is characterized.

$$\int_{x_s^*}^{x_w^*} \ell_w^*(x)dx = \frac{1}{1-\mu_w}\left[c\left(\frac{c^{-1}(1-\mu_w)}{a_h}\right) - c\left(\frac{c^{-1}(1-\mu_s)}{a_h}\right)\right]$$

$$\int_0^{x_s^*} \ell_w^*(x)dx = \frac{c(c^{-1}(1-\mu_s))}{1-\mu_w} - 0 = \frac{1-\mu_s}{1-\mu_w}$$

$$L_w^*(0) = 1 - \frac{1-\mu_s}{1-\mu_w} - \frac{1}{1-\mu_w}\left[c\left(\frac{c^{-1}(1-\mu_w)}{a_h}\right) - c\left(\frac{c^{-1}(1-\mu_s)}{a_h}\right)\right]$$
$$= \frac{\mu_s - \mu_w - \left[c\left(\frac{c^{-1}(1-\mu_w)}{a_h}\right) - c\left(\frac{c^{-1}(1-\mu_s)}{a_h}\right)\right]}{1-\mu_w}$$

The output distributions are

$$L_s^*(x) = \begin{cases} \frac{c(x)}{1-\mu_s}, & 0 \le x \le x_s^* \\ 1, & x_s^* \le x \le x^* \end{cases} \qquad H_s^*(x) = \begin{cases} 0, & 0 \le x \le x_s^* \\ \frac{c(x)}{\mu_s} - \frac{c(x_s^*)}{\mu_s}, & x_s^* \le x \le x_w^* \\ 1 + \frac{c(x/a_h)}{\mu_w} - \frac{c(x^*/a_h)}{\mu_w}, & x_w^* \le x \le x^* \end{cases}$$

$$L_w^*(x) = \begin{cases} \frac{c(x)}{1-\mu_w} + L_w^*(0), & 0 \le x \le x_s^* \\ 1 + \frac{c(x/a_h)}{1-\mu_w} - \frac{c(x_w^*/a_h)}{1-\mu_w}, & x_s^* \le x \le x_w^* \\ 1, & x_w^* \le x \le x^* \end{cases} \qquad H_w^*(x) = \begin{cases} 0, & 0 \le x \le x_w^* \\ 1 + \frac{c(x)}{\mu_w} - \frac{c(x^*)}{\mu_w}, & x_w^* \le x \le x^* \end{cases}$$

where $x_s^* = c^{-1}(1-\mu_s)$, $x_w^* = c^{-1}(1-\mu_w)$, and $x^* = a_h c^{-1}\left(\mu_w + c\left(\frac{c^{-1}(1-\mu_w)}{a_h}\right)\right)$, and $L_w^*(0) = \frac{1}{1-\mu_w}\left[\mu_s - \mu_w - \left[c\left(\frac{c^{-1}(1-\mu_w)}{a_h}\right) - c\left(\frac{c^{-1}(1-\mu_s)}{a_h}\right)\right]\right]$.

If we let $c(e) = ke^\alpha$, with $k > 0$ and $\alpha \ge 1$,[14] then $E[x_s(\mu_s, \mu_w, a^s) + x_w(\mu_w, \mu_s, a^w)]$, the ex-ante expected aggregate output, is

$$\left(\frac{\alpha}{\alpha+1}\right)\left(\frac{1}{k}\right)^{1/\alpha}\left[\left(1 - \frac{1}{a_h^\alpha}\right)\left((1-\mu_w)^{\frac{\alpha+1}{\alpha}} + (1-\mu_s)^{\frac{\alpha+1}{\alpha}}\right) + 2a_h\left(\mu_w + \frac{1-\mu_w}{a_h^\alpha}\right)^{\frac{\alpha+1}{\alpha}}\right].$$

A significant determinant of this total output is μ_w, or the probability the weaker contestant has high ability. Additionally, while an increase in μ_w will increase output, an increase in μ_s will have the opposite effect. For a fixed value of μ_w, an increase in μ_s leads to a reduction in expected aggregate output.

$$\frac{\partial}{\partial \mu_s} E[x_s(\mu_s, \mu_w, a^s) + x_w(\mu_w, \mu_s, a^w)] = -\left(\frac{1}{k}\right)^{1/\alpha}\left(1 - \frac{1}{a_h^\alpha}\right)(1-\mu_s)^{\frac{1}{\alpha}} < 0.$$

This reflects the decrease in overall competition when one contestant is stronger than the other. Therefore, the overall competitiveness of a contest, in the sense of expected output produced, is effected both by the absolute strength of the two contestants and their strengths relative to each other. Intuitively, if a very strong contestant, $\mu_s \approx 1$, and a very weak contestant, $\mu_w \approx 0$, compete against each other, the weak person would likely put in little or no effort, thinking that winning is very unlikely. Additionally, the strong contestant would know the weak contestant is putting in low effort and reduce their effort accordingly.

Repeated contests

We solve for the equilibrium of repeated contests given the parameterization of the cost function, $c(x) = kx^\alpha$, with $\alpha \ge 1$ and $k > 0$.

For the range of $0 \le x < x_{h,*}$ we have $F_1^*(x) = kx^\alpha$, and for the range $x_\ell^* < x \le x_h^*$, we have $F_1^*(x) + v_h = k\frac{x^\alpha}{a_h^\alpha} + k_h$.

For the range $x_{h,*} \le x \le x_\ell^*$, the solution to (†) is

$$F_1^*(x) = Be^{c(x/a_h) - c(x)} + \frac{\frac{\partial}{\partial x}c(x)}{\frac{\partial}{\partial x}c(x) - \frac{\partial}{\partial x}c(x/a_h)}, \text{ with } F_1(x_{h,*}) = kx_{h,*}^\alpha.$$

Solving for, B, the ex-ante distribution function of each contestant on $[x_{h,*}, x_\ell^*]$ is

$$F_1^*(x) = \frac{a_h^\alpha}{a_h^\alpha - 1} - \left(\frac{a_h^\alpha}{a_h^\alpha - 1} - kx_{h,*}^\alpha\right)e^{\frac{k(1-a_h^\alpha)}{a_h{}^\alpha}(x^\alpha - x_{h,*}^\alpha)}.$$

[14] These are implied by the assumptions of a cost function that is strictly increasing and weakly convex

The belief function must satisfy

$$d(\mu(x))\mu'(x) = c'(x) - \frac{1}{a_h}c'(x/a_h) \text{ for } x \in [x_{h,*}, x_\ell^*], \text{ with } \mu(x_{h,*}) = 0.$$

Therefore, on this interval, the belief function is $\mu(x) = k(x^\alpha - x_{h,*}^\alpha)$, and the distribution function can be written as

$$F_1^*(x) = \frac{a_h^\alpha}{a_h^\alpha - 1} - \left(\frac{a_h^\alpha}{a_h^\alpha - 1} - kx_{h,*}^\alpha\right) e^{-\frac{(a_h^\alpha - 1)}{a_h^\alpha}\mu(x)}.$$

Given $F_1^*(x)$ and $\mu(x)$, we calculate the output distribution of the both the high and low ability contestants on $[x_{h,*}, x_\ell^*]$, using $2F_1^*(x) = L_1^*(x) + H_1^*(x)$ and $\mu(x) = \frac{h_1^*(x)}{2f_1^*(x)}$.

$$\begin{aligned}
H_1^*(x) &= H_1^*(x_{h,*}) + 2\int_{x_{h,*}}^{x} \mu(t) f_1^*(t)dt \\
&= 2\mu(t)F_1^*(t)|_{x_{h,*}}^{x} - 2\int_{x_{h,*}}^{x} \mu'(t)F_1^*(t)dt \\
&= 2\left(\frac{a_h^\alpha}{a_h^\alpha - 1} - \left(\frac{a_h^\alpha}{a_h^\alpha - 1} + \mu(x)\right) e^{\frac{1-a_h^\alpha}{a_h^\alpha}\mu(x)}\right)\left(\frac{a_h^\alpha}{a_h^\alpha - 1} - kx_{h,x}^\alpha\right) \\
L_1^*(x) &= 2F_1^*(x) - H_1^*(x) \\
&= \frac{2a_h^\alpha}{a_h^\alpha - 1} + 2\left(\left(\mu(x) + \frac{a_h^\alpha}{a_h^\alpha - 1} - 1\right) e^{\frac{1-a_h^\alpha}{a_h^\alpha}\mu(x)} - \frac{a_h^\alpha}{a_h^\alpha - 1}\right)\left(\frac{a_h^\alpha}{a_h^\alpha - 1} - kx_{h,x}^\alpha\right)
\end{aligned}$$

Let $A = \frac{a_h^\alpha}{a_h^\alpha - 1}$. Ex- ante and ex-interim strategies in the first contest are

$$F_1^*(x) = \begin{cases} kx^\alpha, & 0 \le x \le x_{h,*} \\ A - (A - kx_{h,*}^\alpha)e^{-\mu(x)/A}, & x_{h,*} \le x \le x_\ell^* \\ k(x/a_h)^\alpha + k_h - v_h, & x_\ell^* \le x \le x_h^* \end{cases}$$

$$L_1^*(x) = \begin{cases} 2kx^\alpha, & 0 \le x \le x_{h,*} \\ 2A + 2((\mu(x) + A - 1)e^{-\mu(x)/A} - A)(A - kx_{h,*}^\alpha), & x_{h,*} \le x \le x_\ell^* \\ 1, & x_\ell^* \le x \le x_h^* \end{cases}$$

$$H_1^*(x) = \begin{cases} 0, & 0 \le x \le x_{h,*} \\ 2(A - (A + \mu(x))e^{-\mu(x)/A}(A - kx_{h,*}^\alpha), & x_{h,*} \le x \le x_\ell^* \\ 2k(x/a_h)^\alpha + 2(k_h - v_h) - 1, & x_\ell^* \le x \le x_h^* \end{cases}$$

We use the following conditions to find the unknowns, $x_{h,*}$, x_ℓ^*, x_h^*, v_h and k_h:

1. Continuity of the belief function implies that $\mu(x_\ell^*) = 1$ and $k(x_\ell^{*\alpha} - x_{h,*}^\alpha) = 1$.
2. Since $x_\ell^* = \sup\{BR(a_\ell)\}$, then $L_1^*(x_\ell^*) = 1$.

$$L_1^*(x_\ell^*) = 2A + 2\left((\mu(x_\ell^*) + A - 1)e^{-\mu(x_\ell^*)/A} - A\right)(A - kx_{h,*}^\alpha) = 1$$

$$\Rightarrow kx_{h,*}^\alpha = A - \frac{2A - 1}{2A(1 - e^{-1/A})}, \quad kx_\ell^{*\alpha} = 1 + A - \frac{2A - 1}{2A(1 - e^{-1/A})}$$

3. Continuity of $F_1^*(x)$ at x_ℓ^* gives
$$A - (A - kx_{h,*}^\alpha)e^{-1/A} = \frac{kx_\ell^{*\alpha}}{a_h^\alpha} + k_h - v_h.$$
Substituting from above, the two contest payoff of the high ability contestant is
$$k_h = v_h + \frac{1}{A} + \frac{(2A-1)(1-a_h^\alpha e^{-1/A})}{2Aa_h^\alpha(1-e^{-1/A})}.$$

4. v_h is the expected payoff in the second contest of a contestant with high ability whose ability is known by the opponent in the second contest.
$$v_h = E[v_i(1, \mu(x^{-i}), a_h)] = \frac{1}{A}E[1-\mu(x^{-i})] = \frac{1}{2A}$$
$$\Rightarrow k_h = \frac{3}{2A} + \frac{(2A-1)(1-a_h^\alpha e^{-1/A})}{2Aa_h^\alpha(1-e^{-1/A})}.$$

5. Lastly, $x_h^* = \sup BR(a_h)$ implies $F_1^*(x_h^*) = 1$.
$$F_1^*(x_h^*) = \frac{kx_h^{*\alpha}}{a_h^\alpha} + k_h - v_h = 1$$
$$\Rightarrow kx_h^{*\alpha} = 1 - \frac{(2A-1)(1-a_h^\alpha e^{-1/a})}{2A(1-e^{-1/A})}$$

B Proofs

Lemma 3.1 *In any equilibrium, contestants' distributions of output, $H_s(x), L_s(x), H_w(x)$, and $L_w(x)$, are continuous on $(0, x^*)$, where*
$$x^* \equiv \sup\{BR_s(a_\ell) \cup BR_s(a_h)\} = \sup\{BR_w(a_\ell) \cup BR_w(a_h)\}, \text{ and}$$
$$0 = \inf\{BR_s(a_\ell) \cup BR_s(a_h)\} = \inf\{BR_w(a_\ell) \cup BR_w(a_h)\}$$

Proof. The proof follows in four steps:

1. There is no x at which both contestants both have an atom.
If both contestants played some x with positive probabilities given by p_1 and p_2. Then either contestant can increase output slightly above x, to $x + \varepsilon$. This would increase the payoff of that contestant since the cost of effort is continuous and we can pick ε such that $c(x+\varepsilon) - c(x) < p_2$. However, this implies that x is not a best response of that contestant, a contradiction.

2. If a contestant has an atom, then it is at zero.
Assume that contestant i has an atom at $x > 0$ where x is played is probability $p > 0$. Then by the continuity of the cost function in output, there is a $\delta > 0$ such that $\hat{x} \in (x - \delta/2, x)$, $\hat{x} \notin BR_{-i}(a^{-i})$. This implies, that contestant i would do better by playing $x - \delta/4$, and therefore $x \notin BR_i(a^i)$. This is a contradiction. Therefore the output distribution functions of each type of each contestant is continuous on $(0, \infty)$. This implies that $F_s(x)$, $F_w(x)$

3. If $x > 0$ is not a best response for any ability of one of the contestants, then for all $x' > x$, x' is not a best response for either type of either contestant.
   Step (2) implies that payoffs are continuous, since both the cost function and the probability of winning are continuous in x. Now, since $x \notin \{BR_i(a_\ell) \cup BR_i(a_h)\}$, for some $i = s, w$, $\exists \tilde{x}(a_h), \tilde{x}(a_\ell)$ for which $\pi^i(\tilde{x}(a_h), a_h) > \pi^i(x, a_h) + \varepsilon$ and $\pi^i(\tilde{x}(a_\ell), a_\ell) > \pi^i(x, a_\ell) + \varepsilon$. Then, there is a $\delta > 0$ for which $\pi^i(\tilde{x}(a_h), a_h) > \pi^i(\hat{x}, a_h)$ and $\pi^i(\tilde{x}(a_\ell), a_\ell) > \pi^i(\hat{x}, a_\ell)$, $\forall \hat{x} \in (x, x + \delta)$. Therefore every $\hat{x}$ in this neighborhood cannot be a best response of either type of contestant i. Additionally, no $\hat{x}$ in this neighborhood can be a best response for any type of contestant $-i$, as they could improve utility by lowering output to x. Therefore there is a interval with positive measure for which there is no best responses for either contestant for either type. Let X^* be the set of all outputs that are greater than x and are a best response for some contestant of either type. Let $x_* = \inf\{X^*\}$. Then, necessarily x_* has a gap $(x_* - \delta', x_*]$, $\delta' > 0$ for which there are no best responses. However, since there is an $x \in X^*$ such that $x - x_* < \delta$, x cannot be a best response. Therefore, x_* does not exist and X^* is empty. This implies that $\sup\{BR_s(a_\ell) \cup BR_s(a_h)\} = \sup\{BR_w(a_\ell) \cup BR_w(a_h)\}$. We call this output level x^*.

4. Each contestant has a type who has best response that is arbitrarily close to 0.
   If this were not true, then there is a contestant and an $x > 0$ such that all $\hat{x} \leq x$ are not a best response for any type of that contestant. Then from step (3), that contestant has no best responses. This cannot be true in equilibrium.

□

Corollary 3.3

The ex-interim expected payoff of each contestant is

$$v_s(\mu_s, \mu_w, a_h) = v_w(\mu_w, \mu_s, a_h) = 1 - \mu_w - c\left(\frac{c^{-1}(1 - \mu_w)}{a_h}\right)$$
$$v_s(\mu_s, \mu_w, a_\ell) = \mu_s - \mu_w - \left[c\left(\frac{c^{-1}(1 - \mu_w)}{a_h}\right) - c\left(\frac{c^{-1}(1 - \mu_s)}{a_h}\right)\right]$$
$$v_w(\mu_w, \mu_s, a_\ell) = 0.$$

Proof. Each type of each contestant is indifferent between all outputs in their best response set. Because x^* is in the best response set of high ability contestants, their expected payoffs are equal to the value of winning less the cost of producing output x^*, since, if they produce x^*, they will win with certainty.

$$v_s(\mu_s, \mu_w, a_h) = v_w(\mu_w, \mu_s, a_h) = 1 - c(x^*/a_h) = 1 - \mu_w - c\left(\frac{c^{-1}(1 - \mu_w)}{a_h}\right)$$

Similarly, since $x = 0$ is in the best response set of a low ability contestant, the expected payoffs of low ability contestants is probability they win, given they exert no effort. This is the probability that your opponent puts in no effort.[15]

$$v_s(\mu_s, \mu_w, a_\ell) = (1-\mu_w)L_w(0) = \mu_s - \mu_w - \left[c\left(\frac{c^{-1}(1-\mu_w)}{a_h}\right) - c\left(\frac{c^{-1}(1-\mu_s)}{a_h}\right)\right]$$

$$v_w(\mu_w, \mu_s, a_\ell) = (1-\mu_s)L_s(0) = 0$$

□

Proposition 3.4 (Countervailing incentives)

$\frac{\partial}{\partial \mu_i} E[v_i(\mu_i, \mu_{-i}, a_h)] < 0$ *and* $\frac{\partial}{\partial \mu_i} E[v_i(\mu_i, \mu_{-i}, a_\ell)] > 0$ *for all* $\mu_i \in (0, 1)$.

Proof. In the second contest, for a given pair of beliefs, contestants will expect the following payoffs:

$$v_i(\mu_i, \mu_{-i}, a_h) = 1 - \min\{\mu_i, \mu_{-i}\} - c\left(\frac{c^{-1}(1-\min\{\mu_i, \mu_{-i}\})}{a_h}\right)$$

$$v_i(\mu_i, \mu_{-i}, a_\ell) = \begin{cases} \mu_i - \mu_{-i} - \left[c\left(\frac{c^{-1}(1-\mu_{-i})}{a_h}\right) - c\left(\frac{c^{-1}(1-\mu_i)}{a_h}\right)\right] & \text{if } \mu_i \geq \mu_{-1} \\ 0 & \text{otherwise} \end{cases}$$

For a high ability contestant whose opponent has belief μ_i, the expected payoff in the second contest is given by

$$E[v_i(\mu_i, \mu_{-i}, a_h)] = \int_0^1 \left(1 - \min\{\mu_i, \mu_{-i}\} - c\left(\frac{c^{-1}(1-\min\{\mu_i, \mu_{-i}\})}{a_h}\right)\right) dF_{\mu_{-i}}(\mu_{-i})$$

As the opponent believes the contestant is stronger, the change in expected payoff is

$$\frac{\partial}{\partial \mu_i} E_{\mu_{-i}}[v_i(\mu_i, \mu_{-i}, a_h)] = \left(1 + \frac{\partial}{\partial \mu_i} c\left(\frac{c^{-1}(1-\mu_i)}{a_h}\right)\right)(F_{\mu_{-i}}(\mu_i) - 1)$$

For a low ability contestant, the expected payoff given μ_i is:

$$E[v_i(\mu_i, \mu_{-i}, a_\ell)] = \int_0^{\mu_i} \left(\mu_i - \mu_{-i} + c\left(\frac{c^{-1}(1-\mu_i)}{a_h}\right) - c\left(\frac{c^{-1}(1-\mu_{-i})}{a_h}\right)\right) dF_{\mu_{-i}}(\mu_{-i})$$

The effect of his opponent's beliefs is

$$\frac{\partial}{\partial \mu_i} E[v_i(\mu_i, \mu_{-i}, a_\ell)] = \left(1 + \frac{\partial}{\partial \mu_i} c\left(\frac{c^{-1}(1-\mu_i)}{a_h}\right)\right) F_{\mu_{-i}}(\mu_i)$$

Given the assumptions on the cost of effort, $c'(e) > 0$ and $c''(e) \geq 0$,

$$\frac{\partial}{\partial \mu_i} c\left(\frac{c^{-1}(1-\mu_i)}{a_h}\right) = -\frac{1}{a_h} c'\left(\frac{c^{-1}(1-\mu_i)}{a_h}\right) \frac{1}{c'(c^{-1}(1-\mu_i))} \in \left[-\frac{1}{a_h}, 0\right).$$

[15] Here we technically are assuming the contestant wins all ties at zero, but if the agent exerts a tiny amount of effort and we let that effort shrink to zero, then this is the payoff of the agent in the limit. Since the payoffs are continuous, these limits must be the payoffs of the low ability contestants.

If we define,

$$d(\mu_i) \equiv \left[1 + \frac{\partial}{\partial \mu_i} c\left(\frac{c^{-1}(1-\mu_i)}{a_h}\right)\right] \text{ where } d(\mu_i) \in \left[\frac{a_h - 1}{a_h}, 1\right) \text{ for all } \mu_i,$$

then it is clear that

$$\frac{\partial}{\partial \mu_i} E[v_i(\mu_i, \mu_{-i}, a_h)] = d(\mu_i)(F_{\mu_{-i}}(\mu_i) - 1) < 0$$
$$\frac{\partial}{\partial \mu_i} E[v_i(\mu_i, \mu_{-i}, a_\ell)] = d(\mu_i) F_{\mu_{-i}}(\mu_i) > 0.$$

□

Lemma 4.1 (Monotonic beliefs). *In every SPBE, $\mu(x)$ is weakly increasing in x for all $x \in X_1 = X_1^h \cap X_1^\ell$.*

Proof. Assume otherwise, namely that for a given x and y which are best responses for some ability level we have that $x < y$ and $\mu(x) > \mu(y)$. This implies that $0 \leq \mu(y) < \mu(x) \leq 1$. Then by Bayes' Rule, $h_1(x) > 0$ and $h_1(y) < f_1(y)$, so that $\ell_1(y) > 0$. For the strategies to be optimal it must be that $x \in BR(a_h)$ and $y \in BR(a_\ell)$. Then

$$\begin{aligned}\Pr(\text{win}|y) - c(y) + &E[v_i(\mu(y), \mu(x^{-i}), a_\ell)] \\ &\geq \Pr(\text{win}|x) - c(x) + E[v_i(\mu(x), \mu(x^{-i}), a_\ell)], \text{, and} \\ \Pr(\text{win}|y) - c(y/a_h) + &E[v_i(\mu(y), \mu(x^{-i}), a_h)] \\ &\leq \Pr(\text{win}|x) - c(x/a_h) + E[v_i(\mu_i(x), \mu(x^{-i}), a_h)].\end{aligned}$$

This implies that

$$\begin{aligned}\Pr(\text{win}|y) - \Pr(\text{win}|x) + &E[v_i(\mu(y), \mu(x^{-i}), a_\ell)] - E[v_i(\mu(x), \mu(x^{-i}), a_\ell)] \\ &\geq c(y) - c(x), \text{ and} \\ \Pr(\text{win}|y) - \Pr(\text{win}|x) + &E[v_i(\mu(y), \mu(x^{-i}), a_h)] - E[v(\mu(x), \mu(x^{-i}), a_h)] \\ &\leq c(y/a_h) - c(x/a_h).\end{aligned}$$

The expected payoff in the second contest increases for a low ability contestant as μ increases and for a high ability contestant decreases as μ increases. Then $\mu(x) > \mu(y)$ implies

$$\begin{aligned}E[v_i(\mu(y), \mu(x^{-i}), a_\ell)] - E[v_i(\mu(x), \mu(x^{-i}), a_\ell)] < 0, \\ \text{and } E[v_i(\mu(y), \mu(x^{-i}), a_h)] - E[v(\mu(x), \mu(x^{-i}), a_h)] > 0.\end{aligned}$$

Combine with previous inequalities:

$$c(y) - c(x) < \Pr(\text{win}|y) - \Pr(\text{win}|x) < c(y/a_h) - c(x/a_h)$$

However, since $c''(x) \geq 0$ and $c'(x) > 0$, we must have that

$$\begin{aligned}c(y/a_h) - c(x/a_h) &\leq c(y) - c(y - (y/a_h - x/a_h)) \\ &= c(y) - c\left(\frac{x + (a_h - 1)y}{a_h}\right) < c(y) - c(x).\end{aligned}$$

This is a contradiction. □

Lemma 4.2. *There is no output that is played with positive probability and* $\Pr(win|x) = F_1(x)$ *is continuous.*

Proof. In a symmetric equilibrium, if an output is played with positive probability by one type of contestant, then it must be played with positive probability by this type of both contestants. Let $\hat{x} \in \{X_1^\ell \cup X_1^h\}$ be played with probability $p > 0$. Then

$$\Pr(\text{win}|\hat{x}) + \frac{p}{2} \leq \Pr(\text{win}|x) \text{ for all } x > \hat{x}.$$

Since for some a, $\hat{x} \in BR(a)$, then $\pi(\hat{x}|a) \geq \pi(x|a)$ for all x. This implies that

$$\begin{aligned}\Pr(\text{win}|\hat{x}) - c(\hat{x}/a) + E[v_i(\mu(\hat{x}), \mu(x^{-i}), a^i)] \\ \geq \Pr(\text{win}|x) - c(x/a^i) + E[v_i(\mu(x), \mu(x^{-i}), a^i)]\end{aligned}$$

Combing the above inequalities we have

$$\begin{aligned}\frac{p}{2} &\leq \Pr(\text{win}|x) - \Pr(\text{win}|\hat{x}) \\ &\leq E[v_i(\mu(\hat{x}), \mu(x^{-i}), a^i)] - E[v_i(\mu(x), \mu(x^{-i}), a^i)] + c(x/a^i) - c(\hat{x}/a^i)\end{aligned}$$

By continuity of the cost function, $\exists \varepsilon > 0$ such that for all $x \in (\hat{x}, \hat{x} + \varepsilon)$, we have $c\left(\frac{\hat{x}+\varepsilon}{a^i}\right) - c\left(\frac{\hat{x}}{a^i}\right) < \frac{p}{2}$. Then for each x in this range we know

$$E[v_i(\mu(\hat{x}), \mu(x^{-i}), a^i)] - E[v_i(\mu(x), \mu(x^{-i}), a^i)] > 0.$$

If $a^i = a_\ell$, then $\mu(\hat{x}) > \mu(x)$ and therefore $\hat{x} \in \{X_1^\ell \cap X_1^h\}$. Similarly, if $a^i = a_h$, then $\mu(\hat{x}) < \mu(x)$ and $\hat{x} \in \{X_1^\ell \cap X_1^h\}$. In either case, $\hat{x} \in \{BR(1) \cap BR(a_h)\}$. However, the inequality cannot hold for both $a^i = a_\ell$ and $a^i = a_h$ at the same time, so we have a contradiction. □

We now can use the fact that $F_1(x)$ is continuous in x and we have that $\Pr(\text{win}|x) = \Pr(x < x^{-i}) = \Pr(x \leq x^{-i}) = F_1(x)$. Combined with Lemma 4.1, we have $Pr(\mu(x) < \mu(y)) \leq \Pr(\text{win}|y) = F(y) \leq \Pr(\mu(x) \leq \mu(y))$.

Lemma 4.3 $BR(a_\ell)$ *and* $BR(a_h)$ *are intervals where* $0 = x_{\ell,*} \leq x_{h,*} < x_\ell^* \leq x_h^*$ *and we define* $x_{\ell,*} = \inf\{BR(a_\ell)\}, x_\ell^* = \sup\{BR(a_\ell)\}, x_{h,*} = \inf\{BR(a_h)\}$ *and* $x_h^* = \sup\{BR(a_h)\}$.

Proof. The proof follows in four steps.

1. We first show that $x_{\ell,*} = 0$. We do this by first showing that $x_{\ell,*} \leq x_{h,*}$, and then showing that $x_{\ell,*}$ cannot be larger than zero.

   Let $x_{h,*} < x_{\ell,*}$. Since $x_{h,*} = \inf\{X_1^h\}$, $\forall \varepsilon > 0$, $\exists x_\varepsilon$ such that $x_{h,*} \leq x_\varepsilon < x_{h,*} + \varepsilon$ and $x_\varepsilon \in X_1^h$. In particular, this holds for $\varepsilon^* = x_{\ell,*} - x_{h,*}$. Then $x_{\varepsilon^*} \in \{X_1^h \setminus X_1^\ell\}$ and $\mu(x_{\varepsilon^*}) = 1$. However, from Lemma 4.1 we would have $\mu(x) = 1$ for all $x \in X_1^\ell$, which cannot hold. Therefore $x_{h,*} \geq x_{\ell,*}$.

If $0 < x_{\ell,*} < x_{h,*}$, then let $x_{h,*} - x_{\ell,*} = \delta_1$. Since F_1 is continuous from Lemma 4.2, then $\exists\delta_2$ with $0 < \delta_2 < \delta_1$ such that $\forall x \in (x_{\ell,*}, x_{\ell,*} + \delta_2)$ we have $|F_1(x) - F_1(0)| = |F_1(x) - F_1(x_{\ell,*})| < c(x_{\ell,*}) < c(x_{\delta_2})$. Let $x_{\delta_2} \in X_1^\ell \cap (x_{\ell,*}, x_{\ell,*} + \delta_2)$. Then $\mu(x_{\delta_2}) = 0$ and

$$\begin{aligned} E[\pi^i(0)|a_\ell] &= F_1(0) + E[v_i(\mu(0), \mu(x^{-i}), a_\ell)] \\ &> F_1(x_{\delta_2}) + E[v_i(\mu(x_{\delta_2}), \mu(x^{-i}), a_\ell)] - c(x_{\delta_2}) \\ &= E[\pi^i(x_{\delta_2})|a_\ell] \end{aligned}$$

Then $x_{\delta_2} \notin BR(a_\ell)$, a contradiction.

If $0 < x_{\ell,*} = x_{h,*}$, then $\exists x_\ell, x_h$ such that $x_\ell \leq x_h$, $x_\ell \in X_1^\ell$, $x_h \in X_1^h$, and $F_1(x_\ell) - F_1(x_{\ell,*}) = F_1(x_\ell) < c(x_{\ell,*}) < c(x_\ell)$ and $F_1(x_h) - F_1(x_{h,*}) = F_1(x_h) < c(x_{h,*}/a_h) < c(x_h/a_h)$, by the continuity of F_1.

$x_\ell \in X_1^\ell$ implies that

$$\begin{aligned} E[\pi^i(x_\ell)|a_\ell] &= F_1(x_\ell) - c(x_\ell) + E[v_i(\mu(x_\ell), \mu(x^{-i}), a_\ell)] \\ &\geq F_1(0) - c(0) + E[v_i(\mu(0), \mu(x^{-i}), a_\ell)] = E[\pi^i(0)|a_\ell] \end{aligned}$$

This can hold only if $E[v_i(\mu(x_\ell), \mu(x^{-i}), a_\ell)] > E[v_i(\mu(0), \mu(x^{-i}), a_\ell)]$, which implies that $\mu(x_\ell) > \mu(0)$.

$x_h \in X_1^h$ implies that

$$\begin{aligned} E[\pi^i(x_2)|a_h] &= F_1(x_h) - c(x_h/a_h) + E[v_i(\mu(x_h), \mu(x^{-i}), a_h)] \\ &\geq F_1(0) - c(0) + E[v_i(\mu(0), \mu(x^{-i}), a_h)] = E[\pi^i(0)|a_h] \end{aligned}$$

This can hold only if $E[v_i(\mu(x_h), \mu(x^{-i}), a_h)] > E[v_i(\mu(0), \mu(x^{-i}), a_h)]$, which implies that $\mu(x_h) < \mu(0)$.

Combining these two inequalities leads to $\mu(x_h) < \mu(x_\ell)$. This contradicts Lemma 4.1.

Therefore we must have $0 = x_{\ell,*} \leq x_{h,*}$.

2. We next show that $x_{h,*} \leq x_\ell^*$.

If $x_\ell^* > x_{h,*}$, then $\forall x \in (x_\ell^*, x_{h,*})$, $x \notin \{X_1^\ell \cap X_1^h\}$. Let $x' = \frac{x_\ell^* + x_{h,*}}{2}$ and $\varepsilon = c(x_{h,*}/a_h) - c(x'/a_h)$. There is a $\delta > 0$ such that $\forall x \in (x_{h,*}, x_{h,*} + \delta)$, $F(x) - F(x_{h,*}) < \varepsilon$. Pick an x_δ such that $x_\delta \in (x_{h,*}, x_{h,*} + \delta)$ and $x_\delta \in X_1^h$. Then $F_1(x_\delta) - F_1(x_{h,*}) = F_1(x_\delta) - F_1(x') < \varepsilon$, $c(x_\delta/a_h) - c(x'/a_h) > \varepsilon$, and $E[v_i(\mu(x_\delta), \mu(x^{-i}), a_h)] \leq E[v_i(\mu(x'), \mu(x^{-i}), a_h)]$. Therefore

$$\begin{aligned} E[\pi^i(x')|a_h] &= F_1(x') + E[v_i(\mu(x'), \mu(x^{-i}), a_h)] - c\left(\frac{x'}{a_h}\right) \\ &> F(x_\delta) + E[v_i(\mu(x_\delta), \mu(x^{-i}), a_h)] - c\left(\frac{x_\delta}{a_h}\right) = E[\pi^i(x_\delta)|a_h], \end{aligned}$$

a contradiction. So we can conclude that $x_\ell^* \leq x_{h,*}$.

Also note that we must have $x_\ell^* \leq x_h^*$. If we assume otherwise, then we can find $x \in \{X_1^\ell \setminus X_1^h\}$ where $x > x_h^*$ and $\mu(x) = 0$. Lemma 4.1 rules out this possibility.

We have shown so far that $0 = x_{\ell,*} \leq x_{h,*} \leq x_\ell^* \leq x_h^*$.

3. We next will show that for all $x \in (x_{\ell,*}, x_{h,*})$, $x \in BR(a_\ell)$ and for all $x \in (x_\ell^*, x_h^*)$, $x \in BR(a_h)$.

If $x_{\ell,*} < x_{h,*}$, then let $X_c^\ell = \{x | x \in \{(x_{\ell,*}, x_{h,*}) \setminus BR(a_\ell)\}\}$. If $x \in X_c^\ell$, then $\exists \varepsilon > 0$ such that $E[\pi^i(x)|a_\ell] < E[\pi^i(x')|a_\ell] - \varepsilon$ for all $x' \in \{(x_{\ell,*}, x_{h,*}) \cap X_1^\ell\}$. This implies that:

$$F_1(x) + E[v_i(\mu(x), \mu(x^{-i}), a_\ell)] - c(x) < F_1(x') + E[v_i(\mu(x'), \mu(x^{-i}), a_\ell)] - c(x') - \varepsilon,$$

where $E[v_i(\mu(x), \mu(x^{-i}), a_\ell)] \geq E[v_i(\mu(x'), \mu(x^{-i}), a_\ell)]$ as $\mu(x') = 0$. Therefore $F_1(x) - c(x) < F_1(x') - c(x') - \varepsilon$, and for all $x' > x$ with $x' \in \{(x_{\ell,*}, x_{h,*}) \cap X_1^\ell\}$, $F_1(x') - F_1(x) > c(x') - c(x) - \varepsilon$.

Since F_1 and c are continuous, then there is a $\delta(\varepsilon) > 0$ such that for all $x' \in X_1^\ell$, $|x' - x| \geq \delta(\varepsilon)$. This implies that x is contained in an interval which is a subset of X_c^ℓ. Let a and b be the infimum and supremem of this interval respectively.

- If $b < x_{h,*}$, then $\exists x' < x_{h,x}, x' \in X_1^\ell$ where $|x' - b| < \delta, \forall \delta > 0$. Then, by the continuity of F, $\exists x' \in X_1^\ell$ and $F(x') - F(b) < c(b) - c(\frac{a+b}{2})$. Then we know that

$$F_1(x') - F_1\left(\frac{a+b}{2}\right) < c(b) - c\left(\frac{a+b}{2}\right) \text{ and}$$
$$E[v_i(\mu(x'), \mu(x^{-i}), a_\ell)] \leq E\left[v_i\left(\mu\left(\frac{a+b}{2}\right), \mu(x^{-i}), a_\ell\right)\right].$$

  This implies that $E[\pi^i(x')|a_\ell] < E[\pi^i(\frac{a+b}{2})|a_\ell]$ which contradicts $x' \in BR(a_\ell)$.

- If $b = x_{h,*}$, then $\forall \delta > 0$, $\exists x' \in X_1^h$, s.t. $|x' - b| < \delta$. We again can take $x' \in X_1^h$ such that $F_1(x') - F_1(b) < c(\frac{b}{a_h}) - c(\frac{a+b}{2a_h})$.
  - If $x' \notin X_1^\ell$ then $\mu(x') = 1$, but since $E[v_i(\mu(x'), \mu(x^{-i}), a_h)] \leq E\left[v_i\left(\mu\left(\frac{a+b}{2}\right), \mu(x^{-i}), a_h\right)\right]$, then this contradicts $x' \in BR(a_h)$.
  - If $x' \in X_1^\ell$, then $\mu(x') \in (0, 1)$. If $\mu(x') \leq \mu(\frac{a+b}{2})$, then this contradicts $x' \in BR(a_\ell)$, but if $\mu(x') \geq \mu(\frac{a+b}{2})$, this contradicts $x' \in BR(a_h)$.

Therefore X_c^ℓ must be empty.

If $x_\ell^* < x_h^*$, then let $X_c^h = \{x | x \in \{(x_\ell^*, x_h^*) \setminus BR(a_h)\}$. If $x \in X_c^h$, then $\exists \varepsilon > 0$ such that $E[\pi(x)|a_h] < E[\pi(x')|a_h] - \varepsilon$ for all $x' \in X_1^h$. This implies that

$$F_1(x) - c\left(\frac{x}{a_h}\right) + E[v_i(\mu(x), \mu(x^{-i}), a_h)] < F_1(x') - c\left(\frac{x'}{a_h}\right) + E[v_i(\mu(x'), \mu(x^{-i}), a_h)] - \varepsilon,$$

where $E[v_i(\mu(x), \mu(x^{-i}), a_h)] \geq E[v_i(\mu(x'), \mu(x^{-i}), a_h)]$ as $\mu(x') = 1$. Therefore $F_1(x) - c(\frac{x}{a_h}) < F_1(x') - c(\frac{x'}{a_h}) - \varepsilon$. Since F_1 and c are continuous, then this holds only if $|x' - x| \geq \delta(\varepsilon) > 0$, $\forall x' \in BR(a_h)$. We take a and b to be the infimum and supremum respectively of the interval of X_c^h containing x. Note that $b < x_h^*$, by the definition of x_h^*.

Now, there is an $x' \in X_1^h$ where $|x' - b| < \delta$ for all $\delta > 0$. Therefore there is an $x' \in BR(a_h)$ such that $F_1(x') - F_1(b) < c(\frac{b}{a_h}) - c(\frac{b+a}{2a_h})$. Note that this implies that $F_1(x') - F_1(\frac{b+a}{2}) < c(\frac{x'}{a_h}) - c(\frac{b+a}{2a_h})$. However, this implies that

$$E\left[\pi^i\left(\frac{b+a}{2}\right)|a_h\right] = F_1\left(\frac{b+a}{2}\right) - c\left(\frac{b+a}{2a_h}\right) + E\left[v_i\left(\mu\left(\frac{b+a}{2}\right), \mu(x^{-i}), a_h\right)\right]$$
$$> F_1(x') - c\left(\frac{x'}{a_h}\right) + E[v_i(\mu(x'), \mu(x^{-i}), a_h)] = E[\pi^i(x')|a_h].$$

This contradicts $x' \in BR(a_h)$, and therefore X_c^h must be empty.

4. Lastly, we show that $x_{h,*} < x_\ell^*$, and for all $x \in (x_{h,*}, x_\ell^*)$, $x \in \{BR(a_\ell) \cap BR(a_h)\}$. If $x_\ell^* = x_{h,*}$, then $\forall \delta > 0$, there is $x_\ell \in BR(a_\ell)$ and $x_h \in BR(a_h)$ where $|x_h - x_\ell| < \delta$. Therefore, by the continuity of F_1 and c, there is x_h and x_ℓ for which

$$F_1(x_h) - c\left(\frac{x_h}{a_h}\right) - \left(F_1(x_\ell) - c\left(\frac{x_\ell}{a_h}\right)\right)$$
$$< E[v_i(\mu(x_\ell), \mu(x^{-i}), a_h)] - E[v_i(\mu(x_h), \mu(x^{-i}), a_h)]$$
$$= E[v_i(0, \mu(x^{-i}), a_h)] - E[v_i(1, \mu(x^{-i}), a_h)],$$

since $E[v_i(0, \mu(x^{-i}), a_h)] - E[v_i(1, \mu(x^{-i}), a_h)] > 0$. This implies that

$$E[\pi^i(x_\ell)|a_h] = F_1(x_\ell) - c\left(\frac{x_\ell}{a_h}\right) + E[v_i(\mu(x_\ell), \mu(x^{-i}), a_h)]$$
$$> F_1(x_h) - c\left(\frac{x_h}{a_h}\right) + E[v_i(\mu(x_h), \mu(x^{-i}), a_h)] = E[\pi^i(x_h)|a_h],$$

which cannot be true as $x_h \in BR(a_h)$.

Now define $X_c = \{x | x \in (x_{h,*}, x_\ell^*) \setminus (BR(a_\ell) \cup BR(a_h))\}$. From Lemma 3.1, we know that for all $x' \in \{(x_{h,*}, x_\ell^*) \cap (X_1^\ell \cup X_1^h)\}$, $\mu(x') \in (0,1)$ and therefore $x' \in \{X_1^\ell \cap X_1^h\}$. If $\mu(x') = 1$, we must have $x_\ell^* \leq x'$, a contradiction. Similarly if $\mu(x') = 0$, then we must have $x_{h,*} \geq x'$ which is also a contradiction.

Let $x \in X_c$ be given. Then for all $x', x'' \in \{(x_{h,*}, x_\ell^*) \cap (X_1^\ell \cap X_1^h)\}$ such that $x' < x < x''$ we must have $\mu(x') \leq \mu(x'')$. Let $\mu^* \in [\sup\{\mu(x')\}, \inf\{\mu(x'')\}]$. These are well-defined as there is at least one such x' and x''.

If $\mu(x) \geq \mu^*$ then $E[v_i(\mu(x), \mu(x^{-i}), a_\ell)] \geq E[v_i(\mu(x'), \mu(x^{-i}), a_\ell)]$ for all x' as defined above. Therefore

$$F_1(x') - c(x') + E[v_i(\mu(x'), \mu(x^{-i}), a_\ell)] - \varepsilon_1 > F_1(x) - c(x) + E[v_i(\mu(x), \mu(x^{-i}), a_\ell)]$$

$$\Rightarrow F_1(x') - c(x') - \varepsilon_1 > F_1(x) - c(x)$$

Then, by continuity of F_1 and c, $\exists \delta_1 > 0$ such that $\forall x'$, $|x' - x| > \delta_1$. Then $[x - \delta_1, x] \subset X_c$.

If $\mu(x) < \mu^*$, then $E[v_i(\mu(x), \mu(x^{-i}), a_h)] \geq E[v_i(\mu(x''), \mu(x^{-i}), a_h)]$ for all x'' as defined above. Therefore

$$F_1(x'') - c\left(\frac{x''}{a_h}\right) + E[v_i(\mu(x''), \mu(x^{-i}), a_h)] - \varepsilon_2$$
$$> F_1(x) - c\left(\frac{x}{a_h}\right) + E[v_i(\mu(x), \mu(x^{-i}), a_h)]$$
$$\Rightarrow F_1(x'') - c\left(\frac{x''}{a_h}\right) - \varepsilon_2 > F_1(x) - c\left(\frac{x}{a_h}\right)$$

Then, by continuity, $\exists \delta_2 > 0$ such that $\forall x''$, $|x'' - x| > \delta_2$. Then $[x, x + \delta_2] \subset X_c$.

In either case, if $x \in X_c$, then there is an interval with some supremum b and infimum a such that $x \in (a, b) \subset X_c$.

If $b < x_\ell^*$, then there is an $x' \in \{(x_{h,*}, x_\ell^*) \cap X_1^\ell \cap X_1^h\}$ where $|x' - b| < \delta$ for all $\delta > 0$. Therefore there is an $x' \in \{(x_{h,*}, x_\ell^*) \cap X_1^\ell \cap X_1^h\}$ such that $F(x') - F(b) < c(b/a_h) - c(\frac{b+a}{2a_h})$. Note that this implies that $F_1(x') - F_1(\frac{b+a}{2}) < c(x'/a_h) - c(\frac{b+a}{2a_h})$ and $F_1(x') - F_1(\frac{b+a}{2}) < c(x') - c(\frac{b+a}{2})$.

If $\mu((b + a)/2) < \mu(x')$ then

$$E\left[\pi^i\left(\frac{b+a}{2}\right)|a_h\right] = F_1\left(\frac{b+a}{2}\right) - c\left(\frac{b+a}{2a_h}\right) + E\left[v_i\left(\mu\left(\frac{b+a}{2}\right), \mu(x^{-i}), a_h\right)\right]$$
$$> F_1(x') - c\left(\frac{x'}{a_h}\right) + E[v_i(\mu(x'), \mu(x^{-i}), a_h)] = E[\pi^i(x')|a_h].$$

If $\mu((b + a)/2) \geq \mu(x')$ then

$$E\left[\pi^i\left(\frac{b+a}{2}\right)|a_\ell\right] = F_1\left(\frac{b+a}{2}\right) - c\left(\frac{b+a}{2}\right) + E\left[v_i\left(\mu\left(\frac{b+a}{2}\right), \mu(x^{-i}), a_\ell\right)\right]$$
$$> F_1(x') - c(x') + E[v_i(\mu(x'), \mu(x^{-i}), a_\ell)] = E[\pi^i(x')|a_\ell].$$

In either case, this contradicts $x' \in \{X_1^\ell \cap X_1^h\}$.

If $b = x_\ell^*$, then there is an $x' \in X_1^h$, such that $|x' - b| < \delta$, and $\mu(x') = 1$. This implies that $F_1(x') - F_1(\frac{b+a}{2}) < c(x'/a_h) - c(\frac{b+a}{2a_h})$, and

$$E\left[\pi^i\left(\frac{b+a}{2}\right)|a_h\right] = F_1\left(\frac{b+a}{2}\right) - c\left(\frac{b+a}{2a_h}\right) + E\left[v_i\left(\mu\left(\frac{b+a}{2}\right), \mu(x^{-i}), a_h\right)\right]$$
$$> F_1(x') - c\left(\frac{x'}{a_h}\right) + E[v_i(\mu(x'), \mu(x^{-i}), a_h)] = E[\pi^i(x')|a_h].$$

This contradicts $x' \in X_1^h$. Therefore X_c must be empty and for all $x \in (x_{h,*}, x_\ell^*)$, we must have $x \in \{BR(a_\ell) \cap BR(a_h)\}$.

□

Lemma 4.4 *The belief function and the distribution functions of output, $L_1(x)$ and $H_1(x)$, are continuous in output on $(0, x_h^*)$. Additionally, the belief function is given by $\mu(x) = 0$ for all $x \in [0, x_{h,*}]$, $\mu(x) = 1$ for all $x \in [x_\ell^*, x_h^*]$ and is weakly increasing on $(x_{h,*}, x_\ell^*)$.*

Proof. By definition, distribution functions are right continuous. Lemma 4.2 shows that there no output is played with positive probability by either low or high ability contestants. This implies that the right limit of the distribution function is equal to the left limit at every point. Therefore H_1 and L_1 are continuous and $F_1 = \frac{1}{2}L_1 + \frac{1}{2}H_1$ is also continuous.

To show that $\mu(x)$ is continuous on $(0, x_h^*)$, note that $E[\pi^i(x)|a_\ell]$ is constant for all $x \in BR(a_\ell)$ and $E[\pi^i(x)|a_h]$ is constant for all $x \in BR(a_h)$. Since both $F_1(x)$ and $c(x)$ are continuous on $(0, \infty)$ and $E[v_i(\mu(x), \mu(x^{-i}), a_\ell)] = c(x, 1) - F_1(x) + k_\ell$ on $[0, x_\ell^*]$ for some constant k_ℓ, then $E[v_i(\mu(x), \mu(x^{-i}), a_\ell)]$ must be continuous on this interval. Also, $E[v_i(\mu(x), \mu(x^{-i}), a_h)] = c(\frac{x}{a_h}) - F_1(x) + k_h$ on $[x_{h,*}, x_h^*]$ for some constant k_h, then $E[v_i(\mu(x), \mu(x^{-i}), a_h)]$ is continuous on this interval. Since $E[v_i(\mu(x), \mu(x^{-i}), a_h)]$ is strictly decreasing in $\mu(x)$, and $E[v_i(\mu(x), \mu(x^{-i}), a_\ell)]$ is strictly increasing in $\mu(x)$, then $\mu(x)$ must also be continuous on $BR(a_\ell) \cup BR(a_h) = [0, x_h^*]$.

Using the above, we now show that the set $[0, x_h^*] \setminus X_1$ has no interior, i.e. there can be no interval $[a, b] \subset [0, x_h^*]$ where for all $x \in [a, b]$, $x \notin X_1$. This implies that X_1 is dense in $[0, x_h^*]$.

If we let $[\tilde{a}, \tilde{b}] \subset [0, x_h^*] \setminus X_1$ be given, then define a and b to be the infimum and supremum respectively of the interval in $[0, x_h^*] \setminus X_1$ which contains $[\tilde{a}, \tilde{b}]$. Neither $x_{h,*}$ nor x_ℓ^* can be contained in the interval as they are the limit point of a subset of X_1. Then the interval $[a, b]$ must be contained within either $[0, x_{h,*}]$, $[x_{h,*}, x_\ell^*]$, or $[x_\ell^*, x_h^*]$.

1. If $[a, b] \subset [0, x_{h,*}]$, then for all $x \in [a, b]$, $f_1(x) = 0$ which implies that $F_1(x) = F_1(a)$ and $x \in BR(a_\ell)$. Therefore,

$$E[v_i(\mu(b), \mu(x^{-i}), a_\ell)] - c(b) = E[v_i(\mu(a), \mu(x^{-i}), a_\ell)] - c(a),$$

and $\mu(b) > \mu(a)$. Then for all $\delta > 0$, there is an $x \in X_1$ such that $|x - b| < \delta$. If $x \in X_1^\ell$, we must have $\mu(x) = 0$ or $x \in X_1^h$. Since $\mu(x)$ is continuous, then $\mu(x) \neq 0$, so $x \in X_1^h$. Also if $x \in X_1^\ell$, then $x \in X_1^h$. Either way, for all $\delta > 0$, there must be an $x \in X_1^h$ for which $|x - b| < \delta$. If $x \in X_1^h \setminus X_1^\ell$, then $\mu(x)$=1, and $E[\pi^i(\frac{a+b}{2})|a_h] > E[\pi^i(x)|a_h]$, a contradiction. If $x \in X_1^h \cap X_1^\ell$ then either $E[\pi^i(\frac{a+b}{2})|a_\ell] > E[\pi^i(a)|a_\ell]$ or $E[\pi^i(\frac{a+b}{2})|a_h] > E[\pi^i(x)|a_h]$, again a contradiction. Therefore $[a, b] \not\subset [0, x_{h,*}]$.

2. If $[a, b] \subset [x_{h,*}, x_\ell^*]$, then for all $x \in [a, b]$, $x \in \{BR(a_\ell) \cap BR(a_h)\}$ which implies

$$E[v_i(\mu(b), \mu(x^{-i}), a_\ell)] - c(b) = E[v_i(\mu(a), \mu(x^{-i}), a_\ell)] - c(a),$$
$$E[v_i(\mu(b), \mu(x^{-i}), a_h)] - c(b/a_h) = E[v_i(\mu(a), \mu(x^{-i}), a_h)] - c(a/a_h)$$

This gives

$$E[v_i(\mu(b), \mu(x^{-i}), a_\ell)] - E[v_i(\mu(a), \mu(x^{-i}), a - \ell)] = c(b) - c(a) > 0,$$
$$E[v_i(\mu(b), \mu(x^{-i}), a_h)] - E[v_i(\mu(a), \mu(x^{-i}), a_h)] = c(b/a_h) - c(a/a_h) > 0.$$

However, these inequalities cannot hold at the same time, so $[a,b] \not\subset [x_{h,*}, x_\ell^*]$.

3. If $[a,b] \subset [x_\ell^*, x_h^*]$, then for all $x \in [a,b]$, $x \in BR(a_h)$ and therefore,

$$E[v_i(\mu(b), \mu(x^{-i}), a_h)] - c(b/a_h) = E[v_i(\mu(a), \mu(x^{-i}), a_h)] - c(a/a_h),$$

and $\mu(b) < \mu(a) \leq 1$. Then for all $\delta > 0$, there is an $x \in X_1^h$ such that $|x-b| < \delta$ and $\mu(x) = 1$. However, this contradicts the continuity of $\mu(x)$. Therefore $[a,b] \not\subset [x_\ell^*, x_h^*]$.

Therefore the interior of $[0, x_h^*] \setminus X_1$ is empty, and X_1 is dense on $[0, x_h^*]$.

Since X_1 is dense on $[0, x_h^*]$ we can now show that $\mu(x) = 0$ for any $x \in [0, x_h^*)$. If $\mu(x) = \varepsilon > 0$, then by the continuity of $\mu(x)$, $\exists \delta > 0$ where $\forall x'$, $|x' - x| < \delta$, $\mu(x) > \varepsilon/2$. However for all $\delta > 0$ there is an $x' \in X_1^\ell \setminus X_1^h$ for which $\mu(x') = 0$, a contradiction. Therefore $\mu(x) = 0$ for all $x \in [0, x_h^*)$. Note that $\mu(x_h^*) = 0$ which follows from a similar argument of continuity form the left. Additionally, $\mu(x) = 1$ for all $x \in [x_\ell^*, x_h^*]$.

Lastly we show that $\mu(x)$ is weakly increasing on $[x_{h,*}, x_\ell^*]$. Let $x, y \in [x_{h,*}, x_\ell^*]$ be such that, $\mu(x) > \mu(y)$ and $x < y$. Then there is an x' and y' arbitrarily close to x and y respectively, where $x', y' \in X_1$ and therefore $\mu(x') \leq \mu(y')$. This is not consistent with $\mu(\cdot)$ being continuous, a contradiction. □

Theorem 4.5 (Uniqueness of equilibrium). *There is a unique symmetric perfect Bayes Nash equilibrium* $\{(L_1^*(x_1), L_2^*(x_2|\mu_i, \mu_{-i})), (H_1^*(x_1), H_2^*(x_2|\mu_i, \mu_{-i}))\}$.

Proof. Previous results show that there are three distinct intervals in each equilibrium. We will show that the endpoints of these intervals and the distribution functions on the intervals are completely determined by the first order conditions of the contestants.

The three intervals we investigate are partitioned by the best response sets of the high and low ability contestants. The first is the set of outputs where only low ability contestants are optimizing: $[0, x_{h,*}) = \{BR(a_\ell) \setminus BR(a_h)\}$. Next is the set of outputs where both high and low ability contestants are optimizing $[x_{h,*}, x_\ell^*] = \{BR(a_\ell) \cap BR(a_h)\}$. Lastly is the set of outputs where only high ability contestants are optimizing: $(x_\ell^*, x_h^*] = \{BR(a_h) \setminus BR(a_\ell)\}$.

Conditions for x being in $BR(a_h)$ and $BR(a_\ell)$ are

$$BR(a_h): \; F_1^*(x) + E[v_i(\mu(x), \mu(x^{-i}), a_h)] - c\left(\frac{x}{a_h}\right) = k_h$$

$$BR(x_\ell): \; F_1^*(x) + E[v_i(\mu(x), \mu(x^{-i}), a_\ell)] - c(x) = k_\ell = 0$$

For the range of $0 \leq x < x_{h,*}$, $E[v_i(\mu(x), \mu(x^{-i}), a_\ell)] = 0$ as $\mu(x) = 0$. Therefore $F_1^*(x) = c(x)$ for all $x \in [0, x_{h,*}]$.

For the range $x_\ell^* < x \leq x_h^*$, $E[v_i(\mu(x), \mu(x^{-i}), a_h)] = E_{x_j}[v_i(1, \mu(x_j), a_h)] \equiv v_h$. Then $F_1^*(x) + v_h = c(x/a_h) + k_h$, for all $x \in [x_\ell^*, x_h^*]$.

For the range $x_{h,*} \leq x \leq x_\ell^*$, $x \in \{X_1^\ell \cup X_1^h\}$ implies $x \in \{X_1^\ell \cap X_1^h\}$. Therefore, both low and high ability contestants are indifferent between all outputs in this range. This indifference condition determines the belief function over this interval by subtracting the condition for X_1^ℓ from the condition for X_1^h.

$$E[v_i(\mu(x),\mu(x^{-i}),a_h)] - E[v_i(\mu(x),\mu(x^{-i}),a_\ell)] = c\left(\frac{x}{a_h}\right) + k_h - c(x)$$

Taking the derivative of each side with respect to output,

$$\frac{\partial}{\partial x}(E[v_i(\mu(x),\mu(x^{-i}),a_h)] - E[v_i(\mu(x),\mu(x^{-i}),a_\ell)]) = \frac{\partial}{\partial x}(c\left(\frac{x}{a_h}\right) - c(x))$$

$$\mu'(x)\left[d(\mu(x))(F_\mu(\mu(x))-1) - d(\mu(x))F_\mu(\mu(x))\right] = \frac{1}{a_h}c'\left(\frac{x}{a_h}\right) - c'(x)$$

$$\mu'(x)d(\mu(x)) = c'(x) - \frac{1}{a_h}c'(x/a_h)$$

Note that on this interval, $\mu'(x) > 0$ and therefore, $F_\mu(\mu(x)) = F_1^*(x)$ for all $x \in (x_{h,*}, x_\ell^*)$.

Taking the derivative of the condition for X_1^ℓ and combining with the previous equality:

$$f_1^*(x) + \mu'(x)d(\mu(x))F_1^*(x) = c'(x)$$

$$f_1^*(x) + \left(c'(x) - \frac{1}{a_h}c'\left(\frac{x}{a_h}\right)\right)F_1^*(x) = c'(x)$$

$$f_1^*(x) = \frac{\partial}{\partial x}c(x)(1-F_1^*(x)) + \frac{\partial}{\partial x}c\left(\frac{x}{a_h}\right)F_1^*(x) \quad (\dagger)$$

From continuity of $F_1^*(x)$, we also have that $F_1^*(x_{h,*}) = c(x_{h,*})$. For a given $x_{h,*}$, using the Picard - Lindelöf Theorem[16], we know that there is a unique solution for $f_1^*(x)$ on $[x_{h,*}, x_\ell^*]$, and therefore $F_1^*(x)$ is determined on this interval. Additionally, given $f_1(x)$ on $[0, x^*]$, the endpoints $x_{h,*}$, x_ℓ^*, and x_h^* can be solved for. With $\mu(x)$ characterized, the equilibrium strategies of high ability and low ability contestants can be calculated.

To see why only one such $x_{h,*}$ can lead to an equilibrium, consider a different initial condition, $F^*(\tilde{x}_{h,*}) = c(\tilde{x}_{h,*})$ where $\tilde{x}_{h,*} > x_{h,*}$ and the associated $\tilde{f}_1(x)$ on $[\tilde{x}_{h,*}, \tilde{x}_\ell^*]$. First note that $L(\tilde{x}_{h,*}) = c(\tilde{x}_{h,*}) > c(x_{h,*}) = L(x_{h,*})$. Also, from (†), for each $x \in [\tilde{x}_{h,*}, \tilde{x}_\ell^*]$, $\tilde{f}_1(x) > f_1(x)$. Lastly, $\tilde{\mu}(\tilde{x}_{h,*}) < \mu(\tilde{x}_{h,*})$. Since $\mu(x) = 1 - \frac{\ell(x)}{2f(x)}$, then for all x where $\tilde{\mu}(x) < \mu(x)$, we must have $\tilde{\ell}(x) > \ell(x)$. Then we have that $\tilde{\ell}(x) > \ell(x)$ and $\tilde{\mu}(x) < \mu(x)$ for every $x \in [\tilde{x}_{h,*}, \tilde{x}_{x,*} + \varepsilon]$. In order to get $\tilde{L}(\tilde{x}_\ell^*) = \tilde{\mu}(\tilde{x}_\ell^*) = 1$, there must be an x such that $\tilde{f}(x) = f(x)$ in $[\tilde{x}_{h,*}, \tilde{x}_\ell^*]$, but this cant be true because $\tilde{f}(x)$ and $f(x)$ are different members of the same family of solutions, and cannot cross. Similarly, there cannot be an equilibrium where $\tilde{x}_{h,*} < x_{h,*}$.

Therefore $F_1^*(x)$ is uniquely characterized on X_1 where $\overline{X}_1 = [0, x_h^*]$. Then $L_1^*(x)$ and $H_1^*(x)$ are uniquely determined on this set. These distributions along with the second period output distributions $L_2^*(x|\mu_i, \mu_{-i})$ and $H_2^*(x|\mu_i, \mu_{-i})$ form the unique symmetric Bayes Nash equilibrium. □

[16] The right hand side of (†) is continuous in x and uniformly Lipshitz continuous in $F_1^*(x)$ on the interval of $[x_{h,*}, x_\ell^*]$, see Lindelöf (1894). Also, due to the properties of the cost function, the distribution function is bound between 0 and 1.

References

Aoyagi, Masaki (2010), "Information feedback in a dynamic tournament." *Games and Economic Behavior*, 70, 242–260.

Avery, Christopher (1998), "Strategic jump bidding in english auctions." *The Review of Economic Studies*, 65, 185–210.

Baye, Michael R, Dan Kovenock, and Casper G De Vries (1993), "Rigging the lobbying process: an application of the all-pay auction." *The American Economic Review*, 83, 289–294.

Bergemann, Dirk and Johannes Horner (2014), "Should auctions be transparent?" *Cowles Foundation Discussion Paper*.

Bonatti, Alessandro, Gonzalo Cisternas, and Juuso Toikka (2016), "Dynamic oligopoly with incomplete information." *The Review of Economic Studies*, forthcoming.

Che, Yeon-Koo and Ian Gale (2003), "Optimal design of research contests." *The American Economic Review*, 93, 646–671.

Ederer, Florian (2010), "Feedback and motivation in dynamic tournaments." *Journal of Economics & Management Strategy*, 19, 733–769.

Fullerton, Richard L and Randolph Preston McAfee (1999), "Auctioning entry into tournaments." *Journal of Political Economy*, 107, 573–605.

Harris, Christopher and John Vickers (1987), "Racing with uncertainty." *The Review of Economic Studies*, 54, 1–21.

Hörner, Johannes and Nicolas Sahuguet (2007), "Costly signalling in auctions." *The Review of Economic Studies*, 74, 173–206.

Konrad, Kai A (2012), "Dynamic contests and the discouragement effect." *Revue d'économie politique*, 122, 233–256.

Konrad, Kai A and Florian Morath (2017), "To deter or to moderate? alliance formation in contests with incomplete information." *Economic Inquiry*, forthcoming.

Lazear, Edward P and Sherwin Rosen (1981), "Rank-order tournaments as optimum labor contracts." *The Journal of Political Economy*, 89, 841–864.

Lindelöf, Ernest (1894), "Sur l'application des méthodes d'approximations successives à l'étude des intégrales réelles des équations différentielles ordinaires." *Journal de Mathématiques Pures et Appliquées*, 117–128.

Mirman, Leonard J, Larry Samuelson, and Amparo Urbano (1993), "Duopoly signal jamming." *Economic Theory*, 3, 129–149.

Moldovanu, Benny and Aner Sela (2006), "Contest architecture." *Journal of Economic Theory*, 126, 70–96.

Münster, Johannes (2009), "Repeated contests with asymmetric information." *Journal of Public Economic Theory*, 11, 89–118.

Ortega Reichert, Armando (2000), "A sequential game with information flow." In *The Economic Theory of Auctions I*, 232–54, Edward Elgar Publishing Limited.

Ridlon, Robert and Jiwoong Shin (2013), "Favoring the winner or loser in repeated contests." *Marketing Science*, 32, 768–785.

Rosen, Sherwin (1986), "Prizes and incentives in elimination tournaments." *The American Economic Review*, 701–715.

Siegel, Ron (2009), "All-pay contests." *Econometrica*, 77, 71–92.

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

Spence, Michael (1973), "Job market signaling." *The Quarterly Journal of Economics*, 87, 355–374.

Terwiesch, Christian and Yi Xu (2008), "Innovation contests, open innovation, and multiagent problem solving." *Management Science*, 54, 1529–1543.

Ye, Lixin (2007), "Indicative bidding and a theory of two-stage auctions." *Games and Economic Behavior*, 58, 181–207.

Zhang, Jun and Ruqu Wang (2009), "The role of information revelation in elimination contests." *The Economic Journal*, 119, 613–641.