Anonymous Competitive Contracts

Pelosse Y.*
GATE, University Lyon 2
E-mail address:pelosse@gate.cnrs.fr

February 2003

Abstract

In this paper we consider a situation in which a principal wants to reward the most productive of two agents competing over $T-$periods while, at the same time eliciting a maximal effort level from them. Dubey and Wu (2000) and Dubey and Haimanko (2000) have shown that these two objectives cannot be satisfied simultaneously. Basically this arises because one of the agents may have a considerable lead over his rival at an interim period. For this reason, these authors advocate the use of a spot-check device so as to provide the correct incentives to the agents. However, such a mechanism suffers from a limitation because of its inability to select the biggest producer of the competition. In this paper, we show that there exists a more general mechanism that achieves the two objectives simultaneously. Our analysis is based on a function that assigns to every possible difference in the agents' output a probabitity of winning the prize. We provide a characterization of this optimal function and study its different properties.

JEL Classification Number: C70, D82, J41.

Key Words: Competitive contracts, Difference-form probability assignment functions, Strategic game.

*This paper has benefited from hepful comments of Paul Pezanis Christou, Guillaume Hearinger, Eloïc Peyrache and Marie-Claire Villeval. I am especially grateful to Nicolas Melissas. Finally, I would like to thank Jordi Brandts for his warm hospitality at the Institut d'Analysis Economica (CSCI, Universitat Autonoma de Barcelona) where part of this work has been carried out. The usual disclaimer applies.

1 Introduction

In this paper we consider a situation in which a principal wants to reward the most productive of two agents competing over T-periods. In such a context, two different mechanisms may be designed by the principal. A natural one for him is to observe the agents' outputs over the *whole* competition and then reward the agent producing the biggest output. Such a mechanism has the advantage that the winner of the competition is the most productive agent; a feature shared by a wide variety of contest situations. However, as Dubey and Haimanko (2002) and Dubey and Wu (2001) (Dubey et al. thereafter) have shown, such a mechanism is generally not feasible for the principal. Roughly, this is because a competition over several periods may create an incentive problem in which one of the agents has no chance to win the prize at an interim period. This naturally arises when one of the agent has a considerable lead over his rival. In this situation, there is clearly no point for both agents to work hard. To avoid this problem, Dubey et al. propose spot-check mechanism, leading the principal to base solely his control on a couple of periods randomly drawn. The advantage of this sampling device lies in its incentive effect; no agent will be keen on giving-up the competition at an interim period. In short, this is because no agent is able to develop a sufficient lead over his rival for a suitably chosen number of sampling periods. However, such a mechanism suffers from a limitation because of its inability to select the biggest producer of the competition. For instance, when the information on the rival's output is perfect, an optimal random sampling device advocates that only *two* observations have to be considered by the principal if the competition lasts for a number of periods such that $T \geq 3$. A direct consequence is that the choice of the winner is no longer linked to the T-periods' stream of outputs. In other words, the choice of the winner depends crucially on the choice of the sample which is itself completly random. In this respect, this makes the random check spot device a rather poor selection mechanism. The purpose of this paper is to seek for an alternative mechanism which provides incentives to work hard for both agents and, *at the same time,* selects the most productive agent as the winner.

Economic analysis abounds of situations where the outcome of a competition is commensurate to the cumulative relative performance of the agents over several periods. These situations range from procurement, political and firm organizations to research tournaments and multi-stage sports competitions. As in Dubey et al. our model focuses on a particular situation.

A major assumption made here is indeed that the principal seeks to induce both agents to put in maximal effort level at each time period of the tournament. This implies that the principal's marginal value of the output produced by the agents offsets the marginal cost of an additionnal unit of effort. To motivate this, let us consider the following example.Think about two firms competing in a research tournament. They have to built a prototype for the government. In this context, the information received by an agent on the progress made by his rival at an interim period will affect his incentives to innovate. This arises because an agent may have or not considerably more progressed than rival, perhaps because of its more advanced R&D in the field. In this environnment, the government aims at maximizing the social value of the prototype as it is a vital high tech weapon for the army, a new equipment for medical purposes, or even the high tech manned flight to Mars. In these context, eliciting maximal effort level from the two agents over the whole tournament is optimal. However, if one of the agent gives-up the competition before the end of the tournament the value of the prototype will be smaller than it would have been if more competition had spurred both firms to work at their full capacity.

The mechanism we propose must be generally costlier for the principal than that proposed by Dubey et al[1]. This is fundamentally because the objectives of the principal are *not the same*. The situation we consider is indeed somewhat different. Here, our principal's objective is not only to implement σ^*as an SE but to reward the *most productive agent* over the T-periods as well. Several reasons may be found in favor of this latter objective. In Dubey et al. the use of the sampling device is actually a way of favoring the *least skilled agent* by increasing significantly his probability of winning. In this respect, the use of such a mechanism may be incompatible with the law's requirement because of its biased character. This is often the case in the procurement of public contracts and research tournament. For instance the US National Academy of Sciences recommends the agencies to experiment contest mechanism in sciences and technology through anonymous rules in the sense that the rules must be seen as" *transparent, fair and unbiased"*

In addition, the biased character contained in a random check spot device, makes the principal unable to select efficiently the best-skilled agent when heterogeneous agents' types are considered. This is in contradiction with

[1]However, we conjecture that this will not be the case when the principal's objective to reward the best skilled agent is explicitly formalized.

the selection efficiency that a tournament must possess as recently pointed out by Hvide and Kristiansen (2003). For instance, in the context of a resarch tournament, this means that the sponsor (i.e. the principal) wishes to allocate funding (i.e. the prize) to the researcher of highest ability as he is most likely to make the required innovation. In other words, the principal may not be indifferent ex ante as to who wins the competition. When the principal uses a random spot-check mechanism, the selection efficiency of the tournament decreases; the lower-type agent may win with a significant probability. Basically this is because the dectection of the best skilled agent is increasing in the sample size. In this respect, Dubey and Haimanko analyze a model in which one of the principal cares about who wins; while he is primary interested in maximizing his aggregate output, he may want to see the *best skilled agent* selected as the winner. The mechanism we propose in this paper is perhaps one of the way to solve the underlying dilemma (i.e. full scrutiny versus incentives) arising when this kind of preferences is considered. This is because our mechanism allows the principal to have a full scrutiny of the agent's output compatible with an incentive device. Though, we do not explicitly enter the selection objective the principal's utility, we show that our mechanism proves to be more satisfying than the sampling method in this respect
Finally, let us remark that biasing the tournament toward the weak agents will further deserve the role of selection of a tournament. As pointed out above, contracts based on a spot-check favors the less skilled agents. Then, we may conjecture that such contracts will generally not attract the best skilled agents but rather the pool of agents of a lower type.

Another aspect of our model lies in that we confine our analysis to the class of anonymous competitive contracts. This is not a restrictive assumption as many contracts are based upon this principle. There indeed exists a wide variety of contest situations but an important common principle undelying all these situations is that the rules of the contests are to be based on non-discriminatory or anonymous rules of selection. This implies that the selection and award criteria are not pertained to the type or ability of the agents but they are solely based on the outputs produced by the agents. The feature of such a contract may be justified on some efficiency grounds through the reinforcement of competition among agents. This is generally consistent with most of the designers' objectives to have the capable agent self identify through an adequately calibrated rewards associated to the predetermined objectives.

As said earlier, the current paper analyzes a T -period tournament model

in the same vein as Dubey et al. More precisely, these authors seek to implement a maximal effort level as a strategic equilibrium (SE thereafter) over a T -period tournament game through a random check spot device.Their main contribution is to show that a sampling device may elicit a maximal effort level at each period (σ^* thereafter).We take-up the building block of Dubey and Wu's model but our objective is somewhat different. Our goal, is indeed to show that a certain class of mechanisms implements σ^*as a strategic equilibrium (SE thereafter) and at *the same time* allows the principal to award the prize to the most productive agent. These two objectives are difficult to reconcile without a more general formulation of contracts[2] than those usually defined in the literature.
For instance, let us consider a situation in which two agents compete in a tournament in which they perfectly observe the production made by the rival as the tournament unfolds over time. If they are T observations corresponding to the T periods of the competition, then the following scenario may happen. Namely, it may be that at a time period one of the agent has lagged behind his rival so that he will get the prize or bonus B with probability one. In that situation, neither the best nor the weakest agent would put in maximal effort level for the remaining periods of the competition. This is so because they are respectively in a **good** and **bad scenario** as defined by Dubey et al. In this situation, triggering competition amongst heteregenous agents to a maximal effort level may lead the principal to prefer another mechanism with more flexibility to award the prize. Here we advocate that this flexibility may be found by a noisy observation in the *agents' output difference.*

The building block of our mechanism is indeed determined by what we call a **probability assignment function**. As an example of such a function, consider that used implicitly defined by Dubey et al. With their *assumption*, an agent is awarded of the prize when his performance *through the sample* exceeds his rivals', even by a slim margin; in other words, an agent wins with certainty whenever he gets just one unit of output than his rival through the periods of control. In our model, their mechanism would formally be translated as a function symmetric with respect to 0 and jumping from 0 to one. This assumption about the probability assignment function rules out the possibility of employing a different mechanism that may perform better in the situation we consider. Basically, this is due to the bad and good scenario that may occur at a time period when the overall

[2] This approach was suggested by Jean-François Mertens in Dubey and Wu (2000).

output over the T periods is considered. We indeed show that implementing σ^* as an SE in that situation requires the probability assignment function to be sensitive in the agents' outputs comparison. More precisely, the relevant class of functions we need to consider are of "*difference form*"[3]. By this, we mean that the probability of allocating the prize to an agent depends on the gap between the overall aggregate output of the two agents. In addition to be of difference form, the probability must be increasing in the gap itself. These features of the class of probability assignment functions taken together make the functions under consideration very flexibles. It is noteworthy that the Dubey et al's function may indeed be approximated by the probability assignment functions belonging to this class. A direct consequence is then pertained to the compatibility of the full observation and the ability of the principal to almost always reward the most productive agent. This is the case as the Dubey et al's form of function always awards the prize to the biggest producer. This approximation argument is repeatedly used in our paper and permits to show that the probability of winning of the best skilled agent is always greater than under a sampling device.
Thus, in our model, it turns out that in addition to the prize B and the number of sampled periods n, a contract is fully defined by specifying the function F_n. Actually, we show that a large class of difference probability assignment functions implements σ^*as an SE for *any* finite horizon T extensive form game Ψ. However, this large class of functions will generally entail too much noise in the allocation of the prize for the biggest producer to be the winner with a sufficiently high probability. What must be found is thus some functions which *mimicks* the function that always allocates the prize to the most productive producer while implementing σ^*. In this paper, we formally characterize the class of these functions and prove that they minimize the prize to be handed out to the winner. This turns out to be the case because minimizing the prize for a given sample size coincides with the fact that the most productive agent must always be the winner. The characterization of this function reveals that making deliberatary some kind of very small "mistakes" when allocating the prize is optimal under the situation we consider.

The organization of the paper is as follows. We first lay out the extensive and the sampling rule. We then states the assumptions relative to the contracts we consider. Our first proposition shows that σ^*may be implemented by a certain class of probability assignment functions while a full scrutiny

[3]This terminology is borrowed from Che and Gale (2000) in a different context.

is carried out by the principal over the T periods (section 5). The class of probability assignment functions which allows this result belongs to what we call the **difference probability assignement functions**. We are then in a position to characterize the optimal difference probability assignment functions which belongs to this class. (section 6). As a direct corollary of this, we show that below a certain sample size, we get back the three step function as the optimal one. The next two sections exploits the fact that the class of the optimal functions is actually an approximation of that used by Dubey et al. In this respect, proposition 2 shows that the most productive agent will be almost always awarded of the prize B while this is not the case when the sampling device is introduced (section 7). The last result shows that, at the equilibrium the probability of winning for the best skilled agent is maximized under our mechanism (section 8). We conclude in section 9. All proofs are relegated to the Appendix.

2 The model

In this section we lay out the extensive form game which is essentially the same game as Dubey and Wu. The only difference lies in the definition of a contract which is slightly more general than usually stated with the introduction of the probability assignment function F_n. In our model, a contract is then formally defined by (Ψ, B, n, F_n) where Ψ denotes the extensive form game, B the bonus or prize awarded to the winner n the sample size and F_n the probability assignment function.

2.1 The extensive form game

Here we present an extensive form game $(\Psi, B, n, F_n)$of imperfect information with finite horizon T. The finite set of agents is denoted by $A = \{1, 2, ..., n\}$. In the sequel we will restrict the analysis to the two agent's case. The tournament game is played as follows. At every period, each agent $\alpha \in \{1, 2\}$ simultaneously choose an effort level in the finite set of possible effort levels, $E_\alpha \equiv \left\{\underline{e}_\alpha, ..., \overline{e}_\alpha\right\}$ where $\underline{e}_\alpha$ stands for the minimum effort level that agent α may put in and $\overline{e}_\alpha \equiv e^*_\alpha$ for the maximal one. This decision induces a chance move which selects randomly a nonegative outputs or signals for each agent α from a finite set denoted by $Q_\alpha \equiv \left\{\underline{q}_\alpha, ..., q^j_\alpha, ..., \overline{q}_\alpha\right\}$. Here $\underline{q}_\alpha \equiv q^0$denotes the minimum output that agent α can get at each period, while $\overline{q}_\alpha \equiv q^{m(\alpha)}$ denotes the largest possible output (the number of productible outputs for α is denoted by $card(Q_\alpha) = m(\alpha)$). As it is standard in the principal-agent literature, the probability distribution over the possible outputs q_α^j is conditional on the effort level put in by agent α in the sense of first order stochastic dominance. Formally,

Assumption 1 (First order stochastic dominance) For every agent $\alpha \in \{1,2\}$ the following property holds,

$$\sum_{q \geq z} p_\alpha^{e_\alpha^*}(q) > \sum_{q \geq z} p_\alpha^{e_\alpha}(q),\ \forall k \in Q_\alpha \diagdown \left\{ \underline{q_\alpha} \right\};\ e_\alpha \in E_\alpha \diagdown \{e^*\}$$

In the rest of the paper we assume that this conditional distibution has full support over the set Q_α. This means that $p_\alpha^{e_\alpha}(q^j) > 0, \forall e_\alpha \in E_\alpha$ and $q_\alpha^j \in Q_\alpha$ for any $\alpha \in \{1,2\}$ so that any output level q_α^j can be obtained at every time period. The agents are then led to the next period $t+1$ at the node $\widetilde{w} \equiv (w, e_\alpha, e_\beta, q_\alpha, q_\beta) \in \Omega(t+1)$. The set of *all* agent's node at period t is denoted by $\Omega(t)$ while the set of all decision nodes in the game tree for the agents is $\Omega \equiv \underset{t \in T}{\cup} \Omega(t)$. A typical element of $\Omega(t)$ will be denoted by $w \in \Omega(t) \subset \Omega$ such that, $w \equiv (e_\alpha(\tau), e_\beta(\tau), q_\alpha(\tau), q_\beta(\tau))_{\tau=1}^{t-1} \in \Omega(t)$.

As in Dubey and Wu, the following assumptions pertained to the information pattern of the agents are necessary for some of our results to hold.

Assumption 2 (Imperfect information over the opponent's effort level) We denote by $w \sim (I_\alpha)\ \widetilde{w}$ if two nodes are in the same information set. Each agent *cannot observe* the effort level put in by his rival over the whole tournament.

If $w \equiv (e_\alpha(\tau), e_\beta(\tau), q_\alpha(\tau), q_\beta(\tau))_{\tau=1}^{t-1}$ and $\widetilde{w} \equiv (e_\alpha(\tau), \widetilde{e}_\beta(\tau), q_\alpha(\tau), q_\beta(\tau))_{\tau=1}^{t-1}$ then $w \sim (I_\alpha)\ \widetilde{w}$.

Assumption 3 (Perfect recall) Let I_α the information partition for agent α over Ω. We postulate,

$$\left\{ \begin{array}{c} w \sim (I_\alpha)\widetilde{w} \\ w \equiv (e_\alpha(\tau), e_\beta(\tau), q_\alpha(\tau), q_\beta(\tau))_{\tau=1}^{t-1} \\ \widetilde{w} \equiv (\widetilde{e}_\alpha(\tau), \widetilde{e}_\beta(\tau), \widetilde{q}_\alpha(\tau), \widetilde{q}_\beta(\tau))_{\tau=1}^{t-1} \end{array} \right\} \Rightarrow$$
$$\left\{ \begin{array}{c} t = \widetilde{t} \\ e_\alpha(\tau) = \widetilde{e}_\beta(\tau), \text{for } 1 \leq \tau \leq t-1 \\ (w \mid t') \sim (I_\alpha)(\widetilde{w} \mid t'),\ \text{for } 1 \leq t' < t-1 \end{array} \right\}$$

Assumption 4 (Invariance of output information on memory of effort) There exist partitions $J_{\alpha}(\widetilde{t})$ of $Q_{\alpha}^{\widetilde{t}-1} \times Q_{\beta}^{\widetilde{t}-1}$ for $1 < \widetilde{t} \leq T$ which characterize I_{α} in the sense that:

$S \in I_\alpha$ implies $\exists t \in T$, $(e_{\beta}(\tau))_{\tau=1}^{t-1} \in E_{\alpha}^{t-1}$, $K \in J_\alpha(t)$ such that:

$S = \left\{(e_\alpha(\tau), e_\beta(\tau), q_\alpha(\tau), q_\beta(\tau))_{\tau=1}^{t-1} : e_\beta(\tau) \in E_\beta, \ (q_\alpha(\tau), q_\beta(\tau))_{\tau=1}^{t-1} \in K\right\}$

The payoffs at the terminal nodes, $\widetilde{w} \in \Omega(t+1)$ are determined by a sampling mechanism as indicated below.

2.2 The sampling rule

As said above, we need to set the model in the context of a principal who can use a spot-check device so as to observe the outputs produced throughout the T periods of the tournament. This device amounts to randomly sampling n periods over the T periods of the competition. We define by C_n the set of possible samples that is, $C_n \equiv \{\zeta \subset T : |\zeta| = n\}$ where $|\zeta|$ denotes the cardinal of each sample ζ. As a result, for a tournament of T periods and a sample of size n there are $(T-n)!n!/T$ possible samples ζ .Every sample $\zeta \in C_n$ is susceptible to be picked with a uniform probability. Hence $p(n) \equiv T!/(T-n)!n!$ is the probability that such a sample is to be drawn by the principal. Note that an important assumption lies in the impossibility for an agent to know *when* he is controlled. This implies that no bayesian revision may be carried out by the agents as the tournament unfolds overtime.

2.3 The Normal Form Game $\Gamma(\Psi, B, n, F_n)$

In this section we define the reduced normal form game of the extensive form described above. For convenience, we keep the notations introduced by Dubey and Wu. As in Dubey and Wu, the analysis of the reduced form game is sufficient for our purpose[4].

Definition 1 *Let $\widetilde{w} = (w, e_\alpha, e_\beta, q_\alpha, q_\beta) \in \Omega(t+1)$ such that w is its predecessor. We say that $\widetilde{w}$ is **irrelevant** for the strategy σ_α if either $w \in \Omega(t)$ is irrelevant for σ_α or if $\sigma_\alpha(w) \neq e_\alpha$. The set of all irrelevant nodes is denoted by $\Omega^{\sigma\alpha} \subset \Omega$. If $\Omega^{\widetilde{\sigma}\alpha} = \Omega^{\sigma\alpha}$ and $\sigma_\alpha(w) = \sigma_\beta(w)$ for all $w \in \Omega \backslash \Omega^{\sigma\alpha}$ then $\sigma_\alpha \approx \widetilde{\sigma}_\alpha$.*

[4]More details on the analysis of the normal form game may be found in Dubey and Wu p317.

In the sequel, we denote by $\sum_\alpha$ the set of all *reduced strategies* for agent α. From the definition above it is possible to restrict the analysis of the game with the set of reduced strategies only. The payoff for agent α writes,

$$\sum_{w\in\Omega(T+1)} p^{\sigma}(w)\left[u_{\alpha}(w)-d_{\alpha}(w)\right]\equiv\Pi_{\alpha}(\sigma_{\alpha},\sigma_{\beta})$$

with $p^{\sigma}(w)=\prod_{t=1}^{\widetilde{t}}p^{e_1}(q_1(t))\,p^{e_2}(q_2(t))$, for $\alpha\in\{1,2\}$being the probability of reaching node w under the profile $\sigma\equiv(\sigma_{\alpha},\sigma_{\beta})$.

Definition 2 *Let* $\Omega\equiv\underset{t\in T}{\cup}\Omega(t)$ *be the set of all decisions nodes for all agents in the game tree.* ***A pure strategy for agent*** $\alpha\in\{1,2\}$ *in the normal form game* $\Gamma(\Psi,B,n,F_n)$ *is the mapping* $\sigma_\alpha:\Omega\rightarrow E_\alpha$.

Definition 3 *Let* $\sigma_\alpha\in\sum_\alpha$ *and* $\sigma_\beta\in\sum_\beta$. *Then* σ_α *is a* ***best response*** *to* σ_β *if* $\Pi_\alpha(\sigma_\alpha,\sigma_\beta)\geq\Pi_\alpha(\widetilde{\sigma}_\alpha,\sigma_\beta)$ *for all* $\widetilde{\sigma}_\alpha\in\sum_\alpha$.

Definition 4 *The pair* $(\sigma_\alpha,\sigma_\beta)$ *is called a* ***strategic equilibrium*** *(SE) of* $\Gamma(\Psi,B,n,F_n,)$ *if* σ_α *is a best response to* σ_β *for any* $\alpha\in\{1,2\}$.

Assumption 5 (Utility function) Each agent is endowed of a strictly monotonic utility function $u_\alpha:\mathbf{R}_+\rightarrow\mathbf{R}$ with the following property,

$$u_\alpha(B)\rightarrow\infty \text{ as } B\rightarrow\infty$$

This indicates that for each agent α the utility of getting the prize B at the end of the tournament is sufficiently valued. Further we will denote by d_α the disutility for effort. This is a function such as, $d_\alpha:(E_\alpha)^T\rightarrow R_+$ for $\alpha\in\{1,2\}$.

Assumption 6 (Disutility function) Let $(E_\alpha)^T$ be the set of the vector of effort level put in by α over the T periods. Let $\left(e^*_\alpha,...,e^*_\alpha\right)$ be the vector of unconditional maximal effort level. Then, we have,

$$d_\alpha\left\{\left(e^*_\alpha,...,e^*_\alpha\right)\right\}>d_\alpha\left(e\right)\,\forall e\in(E_\alpha)^T/\left\{\left(e^*_\alpha,...,e^*_\alpha\right)\right\}$$

2.4 Anonymous Competitive Contracts under General Probability Assignment Functions

In this section we give a somewhat more general formulation of a anonymous competitive contract than that generally assumed in the literature. This approach is a direct consequence of the mechanism we want to design. The reason lies in the fact that we need of a more flexible way to determine the winner of the game without reneging on the anonymity character of the contract.

As a result, our contract is fully determined by the prize B, the sample size n and a probability assignment function denoted by F_n. This will imply that unlike most of the tournament literature we do not assume the form of this function but we look for a characterization instead. Roughly, this function will assign a probability of winning for every possible difference between the two agents' output. We precisely define it as follows,

Definition 5 ***The probability assignement function*** $F_{n\alpha}$ *for a sample of size* n *and any agent* $\alpha \in \{1, 2\}$ *is defined by,*

$$F_{n\alpha} : D = \left[-\overline{\psi}n, \overline{\psi}n\right] \rightarrow [0, 1] \text{ for any } \alpha \in \{1, 2\}.$$

where $\overline{\psi} \equiv \overline{q}_\alpha - \underline{q}_\beta$ with $\overline{q}_\alpha \in Q_\alpha$ and $\underline{q}_\beta \in Q_\beta$. In other words, $\overline{\psi}$ denotes the largest possible "distance" in outputs within one period between the two agents. Consider the following vector $x = (x(1), ..., x(n)) \in Q_\alpha^n$ of outputs for agent α over n periods of observations by the principal. the output vector $y \in Q_\alpha^n$ of agent β is similarly defined.The function $F_{n\alpha}$ maps the "distance" between these two vectors of outputs $(x, y) \in Q_{\alpha+}^n \times Q_{\beta+}^n$ into the probability of winning the prize B for every agent α. For convenience, the function is defined on the *convex hull* of the set of possible differences, $E = \left\{-\overline{\psi}n, ..., \overline{\psi}n\right\}$ and is denoted by $D \equiv conv(E)$ thereafter[5]. As said above, we shall restrict our analysis to the class of contracts that verifies the anonymity property. In other words, the functions defined below are not pertained to any agent's label $\alpha \in \{1, 2\}$. This simply means that the principal cannot favor one agent through setting different rules pertained on the agent's type. As an example of a probability assignment function, we give below the following one used by Dubey et al as it will be used repeatedly used in the rest of the paper,

[5] For our purpose, this does not affect any of our results.

Definition 6 *The function* $F_n : \left[-\overline{\psi}n, \overline{\psi}n\right] \to [0,1]$ *is called the* ***3 steps assignment probability function*** *if,*

$$F_n(\psi_\alpha(x,y)) = \begin{cases} 1 \text{ if } \psi_\alpha(x,y) > 0 \\ 1/2 \text{ if } \psi_\alpha(x,y) = 0 \\ 0 \text{ if } \psi_\alpha(x,y) < 0 \end{cases}$$

where $\psi_\alpha(x,y) = \sum\limits_{\tau\in\zeta} x_\alpha(\tau) - \sum\limits_{\tau\in\zeta} y_\beta(\tau)$ with $(x,y) \in Q^n_{1+} \times Q^n_{2+}$. Thus, for any vector of outputs, denotes the distance (or gap) between agent α and agent β once the n observations of the principal have been carried out.Note that this function is symmetric with respect to 0. Moreover, it does not depend on the agent's names. As said above, we want these two properties to be shared by the class of probability assignment functions we will consider in the rest of this paper. In this respect, we restrict the class of probability assignment functions to meet the following requirements,

Assumption 7 (Anonymous probability assignment function)

$$F_1(\psi_1(x,y)) = F_2(\psi_2(y,x)),\ \forall (x,y) \in Q^n_{1+} \times Q^n_{2+},$$

This anonymity property says that if all that has changed is the names of the agents and not anything (such as concerning their respective probability distributions of production) then the probability they have to win the prize with a similar difference in output should be the same. As we restrict the probability assignement function to be anonymous, in the rest of the paper, we shall denote by F_n the probability assignement function for a sample of size n and omit the subscript $\alpha \in \{1,2\}$.

Assumption 8 (Symmetric probability assignment function)

$$F_n(\psi_\alpha(x,y)) = 1 - F_n(-\psi_\alpha(x,y)),\ \forall (x,y) \in Q^n_{\alpha+} \times Q^n_{\beta+},\ \alpha \in \{1,2\}$$

This assumption means that in case of a tie between the two agents the probability for each to win the prize is $\frac{1}{2}$. As a result, the anonymity character of the contract implies that the function is symmetric with respect to 0. Note that the anonymity and symmetric assumptions together imply that the prize will be awarded with probability 1 at the end of the tournament (This restriction seems to necessary for the contract to be credible). The last assumption made in this section restricts the class of functions in the rest of the paper to be monotonic[6].

Assumption 9 (Monotonocity)
F_n belongs to the class of monotonic functions,

If $-n\overline{\psi} < ... < \psi^j < \psi^{j-i} < ... < n\overline{\psi}$ *then* $F_n(\psi_\alpha(x,y) = \psi^j) \leq F_n(\psi_\alpha(x,y) = \psi^{j-i})$ (monotonically increasing) or $F_n(\psi_\alpha(x,y) = \psi^j) \geq F_n(\psi_\alpha(x,y) = \psi^{j-i}))$ (monotonically decreasing), $\forall(x,y) \in Q^n_{\alpha+} \times Q^n_{\beta+}$.

Given the assumptions made above and our results in the rest of the paper, we need to introduce some notations that will be useful in the rest of the paper.

Definition 7 *Let* $1 \leq n \leq T$.

$$\Im_n \equiv \left\{ \begin{array}{c} F_n : D = \left[-\overline{\psi}n, \overline{\psi}n\right] \rightarrow [0,1] \\ \text{monotonically increasing symmetric function on } D \end{array} \right\}$$

As a result, $\Im_n$ is the set of *monotonically increasing symmetric* probability assignment functions for a sample of size n. Note that this set comprises the class of probability assignment functions used by Dubey and Wu defined above. Pending the definition of a *feasible partition* $\widetilde{P}_n$ of D for a sample of size n and a finite horizon T in the Appendix, the set denoted by $\widetilde{\Im}_n$ is defined in the following manner, for any sample size $1 \leq n \leq T$.

Definition 8 $\widetilde{\Im}_n \equiv \left\{ \begin{array}{c} F_n : D = \left[-\overline{\psi}n, \overline{\psi}n\right] \rightarrow [0,1] \\ \textit{symmetric with respect to 0} \\ \textit{and } \text{strictly } \textit{monotonic increasing on } \widetilde{P}_n \subseteq \widehat{P}_n \textit{of } D \end{array} \right\}$

2.5 SE-feasible Sample Size with Difference-Form Probability Assignment Functions

We are now in a position to state a general result about the implementation of σ^* for any sample size n. We indeed show that for a certain class of probability assignment functions (we call it the difference probability assignment function thereafter), it is always possible to implement the maximal effort level profil strategy $\sigma^* = (\sigma_1^*, \sigma_2^*)$ for any sample size $1 \leq n \leq T$. Note that we state this result for *any* sample size because we want to show below that a full observation by the principal will maximize the probability of winning of the *best skilled* agent (not necesserally the most productive) when the

[6] We do not know if this assumption may be relaxed for some of our results.

equilibrium profile σ^* is implemented. Notice that for each agent we are looking for the minimum prize that implements σ^* as a n such that,

$$B_\alpha(\Psi, n, F_n) = \min\left\{B : \sigma^*_\alpha \text{ is a best reply to } \sigma^*_\beta \text{ in the game } \Gamma\left(\Psi, B, n, F_n\right)\right\} \tag{1}$$

The definition given by Dubey and Wu of an SE-feasible sample size is then,

Definition 9 *Let $B(\Psi, n, F_n) \equiv \inf\{B : \sigma^* \text{is an SE of}\, \Gamma(B, \Psi, n, F_n)\}$. A sample size n is said to be **SE-feasible** for Ψ if $B(\Psi, n, F_n) < \infty$.*

Since the principal wants to induce σ^* as an SE, the bonus B which allows to implement this equilibrium has the following form,

$$B(\Psi, n, F_n) = \max\left\{B_1(\Psi, n, F_n), B_2(\Psi, n, F_n)\right\}.$$

The characterization of the probability assignment function is required to ensure that any good or bad scenario is avoided when the sample size n is beyond a certain upper bound given by the function $g(\Psi)$. This function is defined as follows by Dubey and Wu,

Definition 10 *Let $f_\alpha(w)$ be the function which for each node w gives the largest sample size for agent α to put in maximal effort level when he observes w perfectly. For the extensive form game Ψ, we have $g(\Psi) = \min\limits_{\alpha\in\{1,2\}} \min\limits_{S\in I\alpha}$ $\max\left\{f_\alpha(w) : w \in S\right\}$.*

Our first task before solving the optimization's problem lies in finding the conditions for the probability assignment function F_n to meet the constraint imposed on the finitness of the prize $B(\Psi, n, F_n)$ for a sample $n > g(\Psi)$. The precise definition of $f_\alpha(w)$ is given in the Appendix. We now state the main result of this section which will allow us to write the constraint of the optimization problem faced by the principal.

Theorem 1 *Assume (1)-$\underset{\sim}{(9)}$ then any sample size $T \geq n > g(\Psi)$ is SE feasible if and only if $F_n \in \Im_n$.*

This result is in sharp contrast with Dubey and Wu's result of a reduced scrutiny when there exists a certain level of information between the agents.

This first result indeed asserts that even if the agents are *not* ignorant of each other's output, then for a sample size $T \geq n > g(\Psi)$ where a good or a bad scenario may take place, the principal is still able to implement a maximal effort level as an SE with a *finite* prize B. Notice that this mechanism is still credible for the principal even if his objective is indeed to *always* reward the biggest producer. Once the T observations have been carried out by the principal and the difference in output units between both agents is *publicly known*, the principal proceeds to the following device: every possible difference gives rise to a different random draw. For instance, a bag contains 100 sheets of paper, marked of the agents'name such that the proportion of an agent's name varies with the difference in output. If an agent picks his name up, he wins the prize.

Formally, the plausibility of this result stems from the strict positivity of the function F_n over the partition $\widetilde{P}_n$ since this gives rise to a *strict probability of winning* at every nodes for every agent. Note, however, that this condition is *not sufficient* to implement σ^*as an SE as F_n is required to be *strictly* increasing over a certain partition $\widetilde{P}_n$ of D. Actually, this second requirement is induced by the strict positivity of the probability of winning 's condition itself . If the function does not satisfy this requirement on a suitably fine partition $\widetilde{P}_n$ of D, and if a bad or good scenario occurs, then the conditional probability of winning for an agent will be invariant in the effort level put in at some nodes of an information set. This will be the case because in that situation, reducing the "gap" by $\overline{q}_\alpha$ or $\underline{q}_\alpha$ or any intermediate productible output q_α^j induces the same probability of winning for agent α (i.e. $F_n(\psi_\alpha^j(x,y)) = F_n(\psi_\alpha^{j-k}(x,y)$ or $F_n(\psi_\alpha^j(x,y)) = F_n(\psi_\alpha^{j+k}(x,y)$ for some $k = 1, ..., m(\alpha)$ at these nodes so that putting in $\underline{e}_\alpha$ or $\overline{e}_\alpha$ leaves the probability for α to win at these nodes unchanged). This formally arises if the function is not *strictly monotonic increasing* over a suitable partition $\widetilde{P}_n$ of D which we precisely define in the Appendix.

2.6 Characterization of the Class of Optimal Difference Probability Assignment Functions

This section is devoted to the characterization of the class $\Im_n^*$ of optimal difference probability assignment functions F_n^*. By this we mean that the prize $B_\alpha(\Psi, n, F_n)$ must be minimized by the function F_n for each $\alpha \in \{1, 2\}$. It follows, taking $\max\{B_1(\Psi, n, F_n), B_2(\Psi, n, F_n)\}$ that the constraints will be satisfied for both agents. Our objective is indeed to have T as an SE feasible "sample size[7]". For that matter, our objective is now to choose a difference probability assignment function F_T which minimizes the finite prize B such that σ^* is an SE of Γ. In proposition 1 above we have shown that the difference probability assignment functions which allows the implementation of σ^* when $T \geq n > g(\Psi)$ are required to belong to the set $\widetilde{\Im}_n$. As a result, an optimal difference probability assignment functions will have to be in the set $\widetilde{\Im}_n$. The optimal difference probability assignment function must solve the following program:

$$\min_{F_n \in cl(\widetilde{\Im}_n)} \widetilde{B}_\alpha(\Psi, n, F_n) \; \forall \alpha \in \{1,2\}$$

where $\widetilde{B}_\alpha(\Psi, n, F_n) = \left\{ B : \sigma^*_\alpha \text{ is a best reply to } \sigma^*_\beta \text{ in the game } \Gamma(\Psi, B, n, F_n) \right\}$.

In the rest of the paper we call it **program 1**. Note that program 1 is a well defined program as $cl(\widetilde{\Im}_n)$ is by definition a closed set. This implies that for $n > g(\Psi)$ the optimal probability assignment function will have to lie in the *interior* of $cl(\widetilde{\Im}_n)$ for the constraints to be met. The next result gives the precise characterization of the class of the *optimal* difference probability assignment functions F^*_n so that any sample size $T \geq n > g(\Psi)$ be SE feasible.

2.6.1 An Optimal Policy with Imperfect Scrutiny

In this section, we characterize the solution of program 1. As noted earlier, solving program 1 if the probability assignment function is in the closure of $\widetilde{\Im}_n$ amounts to ignoring some of the constraints. In this respect, the next result below charcterize the solution of program 1 by requiring the solution to respect all the constraints so that the functions $\widetilde{F^*}_n$ belong to the interior of $cl(\widetilde{\Im}_n)$. Let us stress that for the next result to hold we *only* need to assume that the principal's objective is to implement σ^* as an SE (as in Dubey and Wu) and he must not be keen on maximizing the probability of winning of the most productive agent.

Proposition 1 *Let (1)-(9) hold and* $\psi_\alpha(x, y) = \sum_{\tau \in \zeta} (x_\alpha(\tau) - y_\beta(\tau))$ *with* $(x, y) \in Q^n_{\alpha+} \times Q^n_{\beta+}$. *Let* $\widetilde{\Im}^*_n$ *denote the class of probability assignment functions that approximate the 3 steps assignment probability function uniformly on* D. *The optimal difference probability assignment function solving program* 1 belongs to $\widetilde{\Im}^*_n$ *and is of the form,*

[7] Though this is not a sample by definition.

$$\widetilde{F}_n^*(\psi_\alpha(x,y)) = \begin{cases} 1 \text{ if } \psi_\alpha(x,y) \geq \overline{\psi}n \\ 1 - c_1\varepsilon \text{ if } \psi^1 \leq \psi_\alpha(x,y) < \overline{\psi}n \\ ... \\ 1 - c_j\varepsilon \text{ if } \psi^j \leq \psi_\alpha(x,y) < \psi^{j-1} \\ ... \\ 1/2 \text{ if } \psi_\alpha(x,y) = 0 \\ ... \\ c_j\varepsilon \text{ if } -\psi^{j-1} \leq \psi_\alpha(x,y) < -\psi^j \\ \\ c_1\varepsilon \text{ if } -\overline{\psi}n < \psi_\alpha(x,y) \leq -\psi^1 \\ 0 \text{ if } \psi_\alpha(x,y) \leq -\overline{\psi}n \end{cases}$$

with $\varepsilon > 0$ *arbitrarily small,* $0 < c_1 < c_2 < ... < c_j$ and $T \geq n > g(\Psi)$.

The above proposition states the property that the optimal way of allocating the prize for the principal when the number of observations is above $g(\Psi)$ must be to create a *small* uncertainty in the reward procedure. One possible interpretation, is that the principal's optimal policy will be to make some *small error of measurement* when controlling agents' outputs. As it is stated in the following Corollary, this property is no longer valid when the number of sampling periods are below a certain threshold given by $g(\Psi)$. This result is in line with Dubey et al's result of a reduced scrutiny since the incentive effect at an interim period is preserved in the same way by creating deliberately an uncertainty. This actually amounts to translating this uncertainty brought about by a spot-check mechanism into a "finer uncertainty". In contrast to the sampling device, the noise generated by our mechanism may be tailored arbitrarily small enough to "almost always" reward the most productive agent. While a random spot- check device's power to select the biggest producer decreases significantly its chance. Once again, to understand the nature of the solution found in Proposition 1, note that the assumption that the most productive agent must be the winner is not called upon. Formally, this is because the unique maximizer is the extreme point of $cl(\widetilde{\Im}_n)$. And, since the solution must be in the interior of this set, the form of the optimal probability assignment function will take the form given in proposition 1 which makes the noise arbitrarily small.

The intuition behind the proof of this proposition is as follows. For a given sample size n, the optimal probability assignment function that minimizes the prize B is nothing but the three step function.(i.e. the function used by Dubey et al.). By theorem 1, we know that a sufficient condition for σ^* continues to hold as an SE of Γ is that F_n belongs to $\widetilde{\Im}_n$. This implies that we must find a function that belongs to the set $\widetilde{\Im}_n$ and approximates the three step function F_n. This two requirements are met with the class of functions as defined above.

Corollary 1 *Let (1)-(9) hold. Then, if $n \leq g(\Psi)$, the optimal probability assignment function F_n^* is the 3 step mechanism given by,*

$$F_n^*(\psi_\alpha(x,y)) = \begin{cases} 1 \text{ if } \psi_\alpha(x,y) > 0 \\ 1/2 \text{ if } \psi_\alpha(x,y) = 0 \\ 0 \text{ if } \psi_\alpha(x,y) < 0 \end{cases} \quad (2)$$

2.7 The No Sampling Policy

Given the results obtained in the previous sections, we are now ready to state and prove the main result of the paper.

Proposition 2 *Assume (1)-(9). Let $n = T$ and any information partition I_α. Then under the difference propbability assignment function , the winner of the prize is approximatively always the most productive agent and* σ^* is SE feasible.

Intuitively, the proof of this result relies on the fact that the principal can always find a probability assignement function $\widetilde{F}_T^* \in \widetilde{\Im}_T^*$ which is "close enough" to F^* while still satisfying the requirements that T is SE-feasible. Recall that under the 3 steps probability assignment function $F^* \in \Im_T$,any agent α gets with *probability one* the prize B if he gets $\psi_\alpha(x,y) > 0$. This would imply that *without* any sampling device, the most productive agent will get the prize with certainty. However, this is not the case if a sampling mechanism is introduced (as in Dubey et al.) since some of the outputs produced by the agents are ignored by the principal. Hence, proving that the 3 steps probability assignment function F^* can be approximated[8] by a certain class of functions belonging to $\Im_{Tk}^* \subset \widetilde{\Im}_T^*$ (For convenience, the definition of $\Im_T^*$ is given in the Appendix) shows that the most productive agent is almost always awarded of the bonus.

Before stating the next result, we need of the assumption that one of the agent is better skilled than his rival in the sense that their mean production per period is different.

[8] More precisely, we make use of a uniform approximation of F^* on D.

Assumption 10 *Let μ_α be the mean of production for agent α. We postulate, $\mu_1 > \mu_2$.*

2.8 The Optimal Policy as a Selection Mechanism

Proposition 3 *Let (1)-(10) hold and $n = T$ and any I_α $\forall\alpha \in \{1, 2\}$. Then, there exists an optimal difference probability assignment function $\widetilde{F}^*_{Tk} \in \widetilde{\Im}^*_T$ such that the principal can implements σ^* as an SE with $T \in \arg\max p_1(\sigma^*, n, \widetilde{F}^*_n)$.*

This result shows that the principal may sustain σ^*as an SE with a finite prize while maximizing the probability of winning of the *best skilled* agent. By proposition 2, we know that without sampling,device, the most productive agent is the one that is awarded of the prize B with a probability close to 1. Moreover, the best-skilled agent's probability of winning the prize B is maximized when there is no sampling device. Thus, using the same approximation argument of F^* as in proposition 2, shows that the probability of winning for the best skilled agent is maximized under $\widetilde{F}^*_T \in \widetilde{\Im}^*_T$ (this occurs since the profile $\sigma^*$is then implemented as an SE). It turns out, that in our setting this amounts to maximizing the probability of agent 1 at the equilibrium profile $\sigma^* = (\sigma^*_1, \sigma^*_2)$.

3 Concluding Remarks

In this paper we have provided an alternative mechanism to Dubey et al's random spot-check device. It has been showed that providing incentives for the agents to put in maximal effort level in a multi-period tournament setting is compatible with a reward procedure in which the prize is almost always allocated to the most productive agent. It is also proved that this corresponds to select the best skilled agent as the winner with a significantly high probability. The mechanism we propose relies on a generalization of the competitive contracts as initially defined by Dubey et al. in a dynamic environnment.

Previous work has indeed assumed that the agent's probability of winning the prize is invariant in the outputs'differences between the rivals. We provide a reason for the principal to observe deliberately the ouputs'differences with a small noise: even though the principal's goal is to reward the biggest producer, allocating the prize with a slight amount of uncertainty is an optimal way to give the proper incentives to the agents to work hard.

The mechanism we propose here is a first step toward a more general formulation of competitive contracts. An interesting direction of research is to compare the optimality of our mechanism with that of Dubey and al. when an explicit principal's objective is to reward the best skilled agent.

Appendix.

1-Preliminaries

In the sequel we will use the following notation $\psi_\alpha^j \equiv \psi_\alpha^j(x,y)$ for the j^{th}possible difference between the vector of output x of agent α and the output vector y of agent β with $x = (x(1), ..., x(n))$ and $y = (y(1), ..., y(n))$. We will use the shorter notation ψ^j to denote the generic j^{th} possible difference in output units.

We say that $\widetilde{w} \equiv (w, e_1, e_2, q_1, q_2) \in \Omega(t+1)$ is an immediate follower of w if $w \in \Omega(t)$. In general, $\widetilde{w}$ is a follower of w if there is a sequence $w_1, w_{2,...,} w_k$ of nodes in $\overline{\Omega} \equiv \overset{T+1}{\underset{t=1}{\cup}} \Omega(t)$ such that $w = w_1$, $\widetilde{w} = w_k$ and each w_{l+1} is an immediate follower of w_l for $1 \leq l \leq k-1$. Denote by $S(w) = \left\{\widetilde{w} \in \overline{\Omega} : \widetilde{w} \text{ is a follower of } w\right\}$.

Define $w \equiv (e_\alpha(\tau), e_\beta(\tau), q_\alpha(\tau), q_\beta(\tau))_{\tau=1}^{t-1} \in \Omega(t) \subset \Omega$.We shall consider the list of all possible outputs of both agents across time that are consistent with w and that leave $\alpha' s$ output at w unspecified by,

$$Q(\alpha, w) = \left\{ \begin{array}{c} ((x_\alpha(\tau))_{\tau \in T \backslash \{t\}}, ((y_\beta(\tau))_{\tau \in T} \in Z_\alpha^{T \backslash \{t\}} \times Z_\alpha^T : (x_\alpha(\tau), (y_\beta(\tau))_{\tau=1}^{t-1} = \\ (q_\alpha(\tau), q_\beta(\tau))_{\tau=1}^{t-1} \end{array} \right\}.$$

Let $\delta(x, y, \zeta) = \sum_{\tau \in \zeta} y(\tau) - \sum_{\tau \in \zeta\ /\{t\}} x(\tau)$. Denote by $\psi_\alpha^j(x,y)$ the distance between α and β at a terminal node $w \in \Omega(T+1)$. We have, $\sum_{\tau \in \zeta} y(\tau) - \sum_{\tau \in \zeta\ /\{t\}} x(\tau) - q_\alpha(t) = \psi_\alpha^j(x,y) \Leftrightarrow q_\alpha(t) = \psi_\alpha^j(x,y) - (\sum_{\tau \in \zeta} y(\tau) - \sum_{\tau \in \zeta/\{t\}} x(\tau))$.

The largest sample size which still incentivizes an agent α to put in maximal effort level is now precisely defined as follows by Dubey and Wu,

Definition 11 *(Dubey and Wu) Let T be the horizon of the game, n the sample size and j any integer such that $-\infty \leq j \leq t-1$. We define $f_\alpha(w) =$*

$\min\left\{\underline{f}_{\alpha}(w), \overline{f}_{\alpha}(w)\right\}$ *to be the largest sample size which still incentivizes* α *at* $w \in \Omega$ *to put in* e^{*}_{α} *when he knows perfectly his rival's output history. Let* Z *denotes the integers. Then,*

$$\underline{f}_{\alpha}(w) = \max\left\{\begin{array}{c} k \in Z : 1 \leq k \leq T, (T-t+1)\left(\overline{q}\alpha - \underline{q}\beta\right) + \\ \overline{h}(k-T+t-1, w) \geq 0 \end{array}\right\}$$

$$\overline{f}_{\alpha}(w) = \max\left\{\begin{array}{c} k \in Z : 1 \leq k \leq T, (T-t+1)\left(\underline{q}\alpha - \overline{q}\beta\right) + \\ \underline{h}(k-T+t-1, w) \leq 0 \end{array}\right\}$$

And,

$$\overline{h}(k-T+t-1, w) = \left\{\begin{array}{c} \max\left\{\sum_{\tau\in\zeta}(q_{\alpha}(\tau) - q_{\beta}(\tau)) : \widetilde{\zeta} \subset \{1, ..., t-1\}, \left|\widetilde{\zeta}\right| = j\right\} \\ \text{if } j > 0 \\ 0 \text{ if } j \leq 0 \end{array}\right.$$

The function $\underline{h}$ is defined by replacing max by min.

Definition 12 *We say that a* ***bad scenario*** *occurs for agent* α *at* w, *given sample size* n, *the probability assgnment function* F_n *if for* $(x_{\alpha}(\tau) - y_{\beta}(\tau))_{\tau\in T^{*}}$ *consistent with* w, *we have* $\exists t \in \zeta$ *and* $\sum\limits_{\tau\in\zeta\diagdown\{t\}}(x_{\alpha}(\tau) - y_{\beta}(\tau) > \overline{q}_{\beta} - \underline{q}_{\alpha}$, $\zeta \in C_n$.

Definition 13 *We say that a* ***good scenario*** *occurs for agent* α *at* w, *given sample size* n, *the probability assgnment function* F_n *if for* $(x_{\alpha}(\tau) - y_{\beta}(\tau))_{\tau\in T^{*}}$ *consistent with* w, *we have* $\exists t \in \zeta$ *and* $\sum\limits_{\tau\in\zeta\diagdown\{t\}}(x_{\alpha}(\tau) - y_{\beta}(\tau) < \underline{q}_{\beta} - \overline{q}_{\alpha}$, $\zeta \in C_n$.

Definition 14 *Consider two arbitrary maps* $\sigma_{\alpha}, \sigma_{\beta}$ *from* Ω *to* $E_{\alpha}, E_{\beta,}$ *respectively and a function* $F_n \in \Im_n$. *The probability induced by* $(\sigma_{\alpha}, \sigma_{\beta})$ *that the play of the game goes through* w *and* α *wins the the prize under sample* ζ *is given by*

Next we define in our context what we call the least and the most refined partition of $D = \left[-\overline{\psi}n, \overline{\psi}n\right]$. Before, let us recall the definition of a partition of a set.

Definition 15 *Let* $[a, b]$ *be a given intervall (*$D = \left[-\overline{\psi}n, \overline{\psi}n\right]$ *in our problem). By a partition* P *of* $[a, b]$ *we mean a set of points* $x_0, x_1, ..., x_n$ $(\psi^{j}, ..., \psi^{n}$ *in our problem) where* $a = x_0 \leq x_1 \leq ... \leq x_n = b$.

In the sequel we denote by $\widehat{P}_n = E = \left\{-\overline{\psi}n, ..., 0, ..., \overline{\psi}n\right\}$the *most refined partition* of D.

Definition 16 *Let* $(x, y) \in Q(\alpha, w)$*. We say that,*

(a) P_n *is the* minimal partition *for a sample of size* n *over* D *if* P_n *is the smallest set* $P_n \subset \widehat{P}_n$ *such that*

$$p_\alpha(x, y, \sigma^*_\alpha(w), \zeta, F_n) > p_\alpha(x, y, \sigma_\alpha(w), \zeta, F_n) \tag{3}$$

for any $\sigma_\alpha(w) \neq e^*_\alpha$, $\forall w \in \Omega$ *and* $\alpha \in \{1, 2\}$.

(b) All the partitions such that (3) holds at every node $w \in \Omega$ *are denoted by* $\widetilde{P}_n$*. They are said to be* feasible partitions.

A1.

Lemma 1 Let $F_n \in \widetilde{\Im}_n$ with any feasible partition $\widetilde{P}_n \subseteq \widehat{P}_n = E$ of D. Then for any $w \in \Omega$, we have, $p_\alpha(x, y, \sigma^*_\alpha(w), \zeta, F_n) > p_\alpha(x, y, \sigma_\alpha(w), \zeta, F_n)$ for $\sigma_\alpha(w) \neq e^*_\alpha$ and $\alpha \in \{1, 2\}$ with $1 \leq n \leq T$ if and only if for any $(x, y) \in Q(\alpha, w)$ and $w \in \Omega$, F_n has at least one pair $\psi^j(x, y), \psi^{j-i}(x, y) \in D$ such that $F_n(\psi^j(x, y)) > F_n(\psi^{j-i}(x, y))$.

proof

Let $p_\alpha(x, y, \sigma_\alpha(w), \zeta, F_n) = \sum_i F_n(\psi^i(x, y)) p^{e^*_\alpha}_\alpha(q_\alpha = q^i)$ be the probability for α to win the prize conditionnal on the realization of $(x, y) \in Q(\alpha, w)$ and on the days in $\zeta \in C_n$ sampled, the effort level $e_\alpha \in E_\alpha$put in at w and the probability assignment function $F_n \in \widetilde{\Im}_n$. Take the functions in $\widetilde{\Im}_n$ defined strictly monotonic increasing on the *most refined* partition $\widehat{P}_n$ of $D = \left[-\overline{\psi}n, \overline{\psi}n\right]$. The condition for

$$p_\alpha(x, y, \sigma^*_\alpha(w), \zeta) > p_\alpha(x, y, \sigma_\alpha(w), \zeta)$$

to be met with the partition $\widehat{P}_n$ for an arbitrary node $w \in \Omega$ is equivalent to,

$$\textstyle\sum_j F_n(\psi^j(x, y)) p^{e^*_\alpha}_\alpha(q_\alpha = q^j) > \sum_j F_n(\psi^j(x, y)) p^{e_\alpha}_\alpha(q_\alpha = q^j) \text{ for any } e_\alpha \in E_\alpha / \left\{e^*_\alpha\right\}.$$

Rearranging,

$$\textstyle\sum_j F_n(\psi^j(x, y)) \left[p^{e*}(q_\alpha = q^i) - p^e(q_\alpha = q^i)\right] > 0$$

Denote by $\Delta p^{e*}(q^j) \equiv p^{e*}(q_\alpha = q^j) - p^e(q_\alpha = q^j)$. If $card(Q\alpha) = k$, we set $q^k \equiv \overline{q}$. Note that,

$$F_n(\psi^k) = F_n(\psi^{k-1}) + \varepsilon_{k-1}, ..., F_n(\psi^{k-(k-1)}) = F_n(\psi^0) + \varepsilon_0 \text{ with } \varepsilon_j > 0 \ \forall j = 1, ..., k.$$

This follows from the strict monotonocity of $F_n$on $\widehat{P}_n$. Now, replace recursivelly each $F_n(\psi^{k-j})$ by its expression, we get,

$F_n(\psi^k) = F_n(\psi^0)+ \sum_{j=0}^{k-1} \varepsilon_j,\ F_n(\psi^{k-1}) = F_n(\psi_0)+ \sum_{j=0}^{k-2} \varepsilon_j, ..., F_n(\psi^0) + \varepsilon_0.$

Hence, we can rewrite (4) as,

$$\left[F_n(\psi^0)+ \sum_{j=0}^{k-1} \varepsilon_j,\right] \Delta p^{e*}(q^k) + \left[F_n(\psi^0)+ \sum_{j=0}^{k-2} \varepsilon_j,\right] \Delta p^{e*}(q^{k-1}) + ... + \left[F_n(\psi^0) + \varepsilon_0\right] > 0.$$

Rearranging,

$F_n(\psi^0)\left[\Delta p^{e*}(q^k) + \Delta p^{e*}(q^{k-1}) + ... + \Delta p^{e*}(q^0)\right] +$

$+\Delta p^{e*}(q^k)\left[\sum_{j=0}^{k-1} \varepsilon_j\right] + \Delta p^{e*}(q^{k-1})\left[\sum_{j=0}^{k-2} \varepsilon_j\right] + ... + \varepsilon_0 \Delta p^{e*}(q^1) > 0$

Note that by definition $\Delta p^{e*}(q^k)+\Delta p^{e*}(q^{k-1})+...+\Delta p^{e*}(q^0) = 0$. Hence, we get

$$\varepsilon_0 \sum_{j=1}^{k} \Delta p^{e*}(q^j) + \varepsilon_1 \sum_{j=2}^{k} \Delta p^{e*}(q^j) + ... + \varepsilon_0 \Delta p^{e*}(q^k) > 0 \qquad (4)$$

We know that $\varepsilon_j > 0$ for $j = 0, ..., k$ since $F_n \in \widetilde{\Im}_n$ and, by the FOSD assumption we have that $\Delta p^{e*}(q^k) = p^{e*}(q_\alpha = q^k) - p^e(q_\alpha = q^k) > 0$ since $q^k \equiv \overline{q}$ and as generally, we have by FOSD assumption that,

$$\sum_{j}^{k} \Delta p^{e*}(q^j) > 0\ \forall j = 1, ..., k.$$

we conclude that (4) holds for any node $w \in \Omega$ if $F_n \in \widetilde{\Im}_n$ with the partition $\widehat{P}_n$. In general, note that (4) holds for any partition $\widetilde{P}_n \subseteq \widehat{P}_n$ such that it exists $\varepsilon_j > 0$ for at least one $j = 0, ..., k$. This implies that (3) holds for any $w \in \Omega$ if $F_n \in \widetilde{\Im}_n$ is defined on a partition $\widetilde{P}_n \subseteq \widehat{P}_n$ such that it exists at least one $\varepsilon_j > 0$ for every $(x, y) \in Q(\alpha, w)$. This implies that the least refined partitions are such that there exists only one $\varepsilon_j > 0$ at every $w \in \Omega$ so that (3) holds at such a w. Clearly if $\varepsilon_j = 0$ for any $j = 0, ..., k$ then (4) does not holds and (3) for at least one node $w \in \Omega$ either. This proves the lemma.

The overall probability of winning for an agent α conditionnal on the strategy profile $(\sigma_\alpha, \sigma_\beta)$,the sample size n and the sample $\zeta \in C_n$is writtten as,

$$p^n_\alpha(\sigma_\alpha, \sigma_\beta, \widetilde{F}_n) = p_\alpha((\sigma_\alpha, \sigma_\beta) \mid n, \widetilde{F}_n w^*)$$

where,
$p((\ \sigma_\alpha, \sigma_\beta) \mid \zeta, \widetilde{F}_n, w) = \sum_{\widetilde{w} \in F(w*) \cap \Omega(T+1)} p^{(\sigma_\alpha \sigma_\beta)}(\widetilde{w})\ \ \widetilde{F}_n(\psi^i(\widetilde{w}, \zeta)).$

When ζ ranges over all samples in C_n are chosen by the principal with uniform probability, this event has probability
$p((\ \sigma_\alpha, \sigma_\beta) \mid n, \widetilde{F_n}, w) = \sum_{\zeta \in C_n} P(n) \sum_{\widetilde{w} \in F(w*) \cap \Omega(T+1)} p^{(\sigma_\alpha \sigma_\beta)}(\widetilde{w})\ \ \widetilde{F}_n(\psi_i(\widetilde{w}, \zeta).$

Hence the overall probability that α wins the prize under $(\sigma_\alpha, \sigma_\beta)$ and a sample of size n is
$p^n_\alpha(\sigma_\alpha, \sigma_\beta, \widetilde{F}_n) = .p_\alpha((\ \sigma_\alpha, \sigma_\beta) \mid n, \widetilde{F}_n, w^*).$

Definition 17 *Fix $w \in \Omega$. Suppose $\widehat{\sigma}\alpha$ and $\widetilde{\sigma}\alpha$ are two maps from Ω to E_α which satisfy:*

$(i)\widehat{\sigma}_\alpha(\widetilde{w}) = \widehat{\sigma}_\alpha(\widetilde{w})$ for $\widetilde{w} \in \Omega \backslash \{w\}; (ii)\widehat{\sigma}_\alpha(\widetilde{w}) = \widehat{\sigma}_\alpha(\widetilde{w}) = e^*_\alpha$ for $\widetilde{w} \in S(w); (iii)\widehat{\sigma}_\alpha(\widetilde{w}) \neq e^*_\alpha = \widehat{\sigma}_\alpha(\widetilde{w})$. Then we write $\widehat{\sigma}_\alpha(\widetilde{w}) \succ_w \widehat{\sigma}_\alpha(\widetilde{w})$.

From Lemma 1, the results of Dubey and Wu follows.

Lemma 2 (Dubey and Wu). *Fix $w \in \Omega(t)$ and $\zeta \subset T$. Suppose $\widehat{\sigma}_\alpha \succ_w \widetilde{\sigma}_\alpha$ and $F_n \in \widetilde{\Im}_n$. Then*

1 $p_\alpha((\ \widehat{\sigma}_\alpha(w), \sigma^*_\beta(w) \mid F_n, \zeta, w) \geq p_\alpha((\widehat{\sigma}_\alpha(w), \sigma^*_\beta(w) \mid F_n, \zeta, w)$

2 $p_\alpha((\ \widehat{\sigma}_\alpha, \sigma^*_\beta \mid F_n, n, w) \geq p_\alpha((\widetilde{\sigma}_\alpha, \sigma^*_\beta \mid F_n, n, w)$

3 $p^n_\alpha((\ \widehat{\sigma}_\alpha, \sigma^*_\beta, F_n) \geq p^n_\alpha((\widetilde{\sigma}_\alpha, \sigma^*_\beta, F_n)$

Proof The proof is similar to Dubey and Wu, and we rephrase it only for convenience when $F_n \in \widetilde{\Im}_n$.

If $t \in \zeta$, we can write, $p_\alpha((\ \sigma_\alpha(w), \sigma^*_\beta(w) \mid \zeta, w) = p^{(\sigma_\alpha, \sigma_\beta)}(w)$

$$\times \left(\sum_{(x,y) \in Q(\alpha,w)} \left(\prod_{\tau=t}^{T} p^{e\beta^*}_\beta\left(y\left(\tau\right)\right) \right) \left(\prod_{\tau=t}^{T} p^{e\beta^*}_\beta\left(y\left(\tau\right)\right) \right) p\alpha(x, y, \sigma\alpha\left(w\right), \zeta, F_n) \right)$$

for any $F_n \in \widetilde{\Im}_n$ and any σ_α. By definition of $\succ_w$above, we know that $p^{(\widehat{\sigma}_\alpha, \sigma^*_\beta)}(w) = p^{(\widetilde{\sigma}_\alpha, \sigma^*_\beta)}(w)$. By Lemma 1, $p_\alpha(x, y, e^*_\alpha, \zeta, F_n) > p_\alpha(x, y, e_\alpha, \zeta, F_n)$ for all $e_\alpha \in E_\alpha / \left\{e^*_\alpha\right\}$, $F_n \in \widetilde{\Im}_n$ proves the strict inequality of (1). And, by a similar argument, when $t \notin \zeta$, there is equality in (1). Inequality (2) follows from (1) and by the definition of $p_\alpha((\ \sigma_\alpha(w), \sigma^*_\beta(w) \mid \zeta, w)$. Equation (3) holds by definition of $p^n_\alpha((\ \widehat{\sigma}_\alpha, \sigma^*_\beta, F_n)$ and (2) and because $p\alpha((\widehat{\sigma}_\alpha, \sigma^*_\beta \mid F_n, n, w) = p_\alpha((\widetilde{\sigma}_\alpha, \sigma^*_\beta \mid F_n, n, w)$ for any $w' \in \Omega(t) \setminus \{w\}$.

Before proving propostion 1, we state the following claim of Dubey and Wu,

Claim 1 *Let* $F_n \in \widetilde{\Im}_n$ *then,* $B(\Psi, n, F_n) < \infty \Leftrightarrow prob^n_\alpha(\sigma^*_\alpha, \sigma^*_\beta, F_n) > prob^n_\alpha((\sigma_\alpha, \sigma^*_\beta, F_n)\ \forall\alpha \in \{1,2\}$ and $\sigma_\alpha \in \sum_\alpha / \{\sigma^*_\alpha\}$.

Proof For a proof, see Dubey and Wu p329.

A2.

Proof of proposition 1 In the sequel we denote by $\psi^j(x,y)$ the j^{th}difference between the two agents and we always consider $(x,y) \in Q(\alpha, w)$.We first state that $prob^n_\alpha(\sigma^*_\alpha, \sigma^*_\beta, F_n) > prob^n_\alpha((\sigma_\alpha, \sigma^*_\beta, F_n)\ \forall\sigma_\alpha \in \sum_\alpha / \{\sigma^*_\alpha\}$and $T \geq n > g(\Psi) \Rightarrow F_n \in \widetilde{\Im}_n$. To show this, we must check that *each* of the constraints required for the class of functions to belong to $\widetilde{\Im}_n$ are necessary conditions.

Let $n > g(\Psi)$, it exists an agent α and an information set $\widetilde{S} \in I_\alpha$ where $n > f(\widetilde{w})\ \forall\ \widetilde{w} \in \widetilde{S}$ and it is either a good scenario (i.e. $n > \overline{f}_\alpha(\widetilde{w})$) or a bad scenario (i.e. $n < \underline{f}_\alpha(\widetilde{w})$).

Let $\widetilde{S}$ correspond to $(\widetilde{e^*_\alpha}(\tau))^{t-1}_{\tau=1}$and $K \in J_\alpha(t)$. Now, consider an information set S that corresponds to $(\widetilde{e^*_\alpha}(\tau))^{t-1}_{\tau=1}$and K.(such an information set exists by the full support assumption.) Some node in S is reached with a positive probability under $(\sigma^*_\alpha, \sigma^*_\beta)$ (by the full support assumption). Since $f(\widetilde{w})$ depends only on the output history leading to w each node in S is either a good or a bad scenario (output history over S and $\widetilde{S}$ yield K). Therefore at all terminal nodes that follow from an arbitrary $w \in S$, α wins or loses under the 3 step mechanism. Let $(x,y) \in Q(\alpha, w)$ and $t \in \zeta$.(if $t \notin \zeta$, then this is trivial).

(i) Let $F_n \notin \widetilde{\Im}_n$. Recall that, by definition, if $F_n \in \widetilde{\Im}_n$ then F_n is strictly monotonic increasing on $\widetilde{P}_n \subseteq \widehat{P}_n \equiv E$. Now, suppose that $F_n \notin \widetilde{\Im}_n$ *only* because this function is strictly monotonic increasing on a non feasible partition (i.e. on a set smaller than the minimal partition). Then, by a direct application of lemma 1 we get $p_\alpha(x, y, \sigma_\alpha(w), \zeta) = p_\alpha(x, y, \sigma^*_\alpha(w), \zeta)$ with $\sigma_\alpha(w) \neq e^*_\alpha$. Consider σ_α such that $\sigma_\alpha(w) \neq e^*_\alpha$ for any $w \in S$ but $\sigma_\alpha(w) = \sigma^*_\alpha(w)$ for any $w \notin S$. By lemma 2, this implies that $prob^n_\alpha(\sigma^*_\alpha\sigma^*_\beta, F_n)) = prob^n_\alpha((\sigma_\alpha\sigma^*_\beta, F_n))\ \forall\sigma_\alpha \in \sum_\alpha / \{\sigma^*_\alpha\}$, by claim 1 we get $B(\Psi, n, F_n) = \infty$ a contradiction.

(ii) Let $F_n \notin \widetilde{\Im}_n$. Let $F_n \notin \widetilde{\Im}_n$ *only* because F_n is a weakly decreasing function on the partition $\widetilde{P}_n$. Consider as before $(x,y) \in Q(\alpha, w)$ and $\zeta \in C_n$ such that $t \in \zeta$. Let $-\psi^i(x,y) = q^i - \delta(x,y,\zeta)$ and $-\psi^{i-j}(x,y) = q^{i-j} - \delta(x,y,\zeta)$ with $-\psi^i(x,y$)and $-\psi^{i-j}(x,y) \in \widetilde{P}_n$.

$a)$ If $F_n(-\psi^i(x,y)) = F_n(-\psi^{i-j}(x,y))$ for any $j = 0, ..., k$, then by (4) of lemma 1 with $\varepsilon_j = 0, \forall j = 0, ...k$, we have, $p_\alpha(x, y, \sigma_\alpha(w), \zeta) =$

$p_\alpha(x, y, \sigma^*_\alpha(w))$ with $\sigma_\alpha(w) \neq e^*_\alpha$. Consider σ_α such that $\sigma_\alpha(w) \neq e^*_\alpha$ for any $w \in S$ but $\sigma_\alpha(w) = \sigma^*_\alpha(w)$ for any $w \notin S$. By lemma 2, this implies that $prob^n_\alpha(\sigma^*_\alpha \sigma^*_\beta, F_n)) = prob^n_\alpha((\sigma_\alpha \sigma^*_\beta, F_n))$ $\forall \sigma_\alpha \in \sum_\alpha / \left\{\sigma^*_\alpha\right\}$, by claim 1 we get $B(\Psi, n, F_n) = \infty$ a contradiction.

b) If $F_n(-\psi^{i-1}(x, y)) > F_n(-\psi^i(x, y))$ with $-\psi^{i-1}(x, y) {<} {-}\psi^i(x, y)$ for some $i = 0, ..., j$. If $\varepsilon_j \leq 0, \forall j = 0, ...k$ in (4) of lemma 1 we conclude that $p(x, y, \sigma^*_\alpha(w), \zeta) \leq p(x, y, \sigma_\alpha(w), \zeta)$. Consider σ_α such that $\sigma_\alpha(w) \neq e^*_\alpha$ for any $w \in S$ but $\sigma_\alpha(w) = \sigma^*_\alpha(w)$ for any $w \notin S$. By lemma 2, this implies that $prob^n_\alpha(\sigma^*_\alpha, \sigma^*_\beta, F_n)) = prob^n_\alpha((\sigma_\alpha, \sigma^*_\beta, F_n))$ $\forall \sigma_\alpha \in \sum_\alpha / \left\{\sigma^*_\alpha\right\}$, by claim 1 we get $B(\Psi, n, F_n) = \infty$ a contradiction.

Next we check that $F_n \in \widetilde{\Im}_n \Rightarrow prob^n_\alpha(\sigma^*_\alpha, \sigma^*_\beta, F_n) > prob^n_\alpha((\sigma_\alpha, \sigma^*_\beta, F_n)$ $\forall \sigma_\alpha \in \sum_\alpha / \left\{\sigma^*_\alpha\right\} \text{for} T \geq n > g(\Psi)$

Suppose that $F_n \in \widetilde{\Im}_n$. By application of lemma 1, we get that, $p(x, y, \sigma^*_\alpha(w), \zeta) > p(x, y, \sigma_\alpha(w), \zeta)$ always holds for $\sigma_\alpha(w) \neq \sigma^*_\alpha(w)$ at node $w \in S$. This implies that

$prob^n_\alpha(\sigma^*_\alpha, \sigma^*_\beta, F_n) > prob^n_\alpha((\sigma_\alpha, \sigma^*_\beta, F_n)$ $\forall \sigma_\alpha \in \sum_\alpha / \left\{\sigma^*_\alpha\right\}$ and by claim 1 that $B(\Psi, n, F_n) < \infty$.

A3.

Proof of proposition 2

Step 1 First, note that program 1 is equivalent to the following one:

$$\max_{F_n \in cl(\Im_n)} \left\{prob^n_\alpha(\sigma^*_\alpha, \sigma^*_\beta, F_n) - prob^n_\alpha(\sigma_\alpha, \sigma^*_\beta, F_n)\right\}$$

with $\sigma_\alpha \in \sum_\alpha / \left\{\sigma^*_\alpha\right\}$, $\alpha \in \{1, 2\}$.

Step 2 We first recall the definition of sequential compactness.

Definition. A topological space X ($\Im_n$ in our problem) is called sequentially compact if every infinite sequence from X has a convergent subsequence.We first show that the set of the symmetric functions denoted by $\Im_n$ is sequentially compact. We know that $\Im_n$ is defined by,

$$\Im_n \equiv \left\{ \begin{array}{c} F_n : D = \left[-\overline{\psi}n, \overline{\psi}n\right] \rightarrow [0, 1] \\ \text{is a symmetric non-decreasing function on } D \end{array} \right\}$$

Applying Helly's Selection Principle to the sequence $\{F_{nk}\}$ shows that there is a *subsequence* of F_{nk}which converges pointwise (i.e. in the topology of pointwise convergence[9]) to F_n for every $\psi^j \in D$. By a theorem (see

[9]This topology is sometimes defined as the open-point topology. For a definition, see Munkres p280.

e.g. Munkres p280) this is equivalent to say that for every $\psi^j \in D$, the subsequence of $F_{nk}(\psi^j)$ converges to $F_{nk}(\psi^j)$. From this it follows that every sequence $\{F_{nk}\}$ in $\Im_n$ has a convergent subsequence proving that $\Im_n$ is sequentially compact.

Step 3.To show that the objective function is continuous (in the topology of pointwise convergence) on $\Im_n$ we note that $\Delta(\widetilde{F}_n)$ is a linear function since we have,

$\Delta_\alpha(\widetilde{F}_n) = \sum_{\zeta \in C_n} P(n) \sum_{\widetilde{w} \in F(w*)\cap\Omega(T+1)} \Delta p^{(\sigma_\alpha \sigma_\beta)}(\widetilde{w})\ \widetilde{F}_n(\psi_i(\widetilde{w},\zeta))$ may be rewritten as $\Delta_\alpha(\widetilde{F}_n) =$

$$\widetilde{F}_n(\psi_i(\widetilde{w_1},\zeta)) \left[\Delta p^{(\sigma_\alpha \sigma_\beta)}(\widetilde{w_1})P(n_1) + ... + \Delta p^{(\sigma_\alpha \sigma_\beta)}(\widetilde{w_1})P(n_l)\right] + ...$$
$$\widetilde{+F}_n(\psi_i(\widetilde{w_k},\zeta)) \left[\Delta p^{(\sigma_\alpha \sigma_\beta)}(\widetilde{w_k})P(n_1) + ... + \Delta p^{(\sigma_\alpha \sigma_\beta)}(\widetilde{w_k})P(n_l)\right].$$

As every linear function is continuous, we have that $\Delta(\widetilde{F}_n)$ is a continuous function in the topology of pointwise convergence.

Step 4. Now we show that that a bounded and continuous function defined on a sequentially compact set has a maximum. We know that $\Delta_\alpha(\widetilde{F}_n)$ is bounded by 1 since it is the difference between two probabilities as noted above. Define

$\Delta^*_\alpha(\widetilde{F}_n) = \sup_{\widetilde{F}_n \in \widetilde{\Im}_n} \Delta_\alpha(\widetilde{F}_n)$ and let $\widetilde{F}_{nk}$ be a sequence in $\widetilde{\Im}_n$ such that

$\Delta_\alpha(\widetilde{F}_{nk}) \geq \Delta^*_\alpha(\widetilde{F}_n) - \dfrac{1}{k}, k \in N$. Since $\widetilde{\Im}_n$ is compact, every sequence has a convergent subsequence. Therefore $\left\{\widetilde{F}_{nk}\right\}_{k\in N}$ has a convergent subsequence $\left\{\widetilde{F}_{nk(1)}\right\}_{k(1)\in N}$ that converges to $\widetilde{F^*}_n$. since $\Delta^*_\alpha(\widetilde{F}_n)$ is continuous at $\widetilde{F^*}_n$, we have that $\lim_{k(1)\to\infty} \Delta_\alpha(\widetilde{F^*}_{nk(1)}) = \Delta^*_\alpha$.

Step 5. Since the maximization problem defined above has a maximum, we can now proceed by characterizing the form of the maximizer. For that matter, note that the objective function is a linear function in the variable $\widetilde{F}_n$ as shown in step 2 above. As a result, the maximiser will be an extreme point of the set $cl(\widetilde{\Im}_n)$. Because the class of relevant functions we investigate must be symmetric with respect to 0, the set of extreme points of $cl(\widetilde{\Im}_n)$is a singleton and is of the form,

$$F^*_n(\psi_\alpha(x,y)) = \begin{cases} 1 \text{ if } \psi_\alpha(x,y) > 0 \\ 1/2 \text{ if } \psi_\alpha(x,y) = 0 \\ 0 \text{ if } \psi_\alpha(x,y) < 0 \end{cases}$$

Step 5 (Uniform approximation of F^*) In the sequel, we denote by $\widetilde{\Im}^*_n$ the class of functions of the form,

$$F^*_{nk}(\psi_\alpha(x,y)) = \begin{cases} 1 \text{ if } \psi_\alpha(x,y) \geq \overline{\psi}n \\ 1-\frac{c_1}{2^k} \text{ if } \psi^1 \leq \psi_\alpha(x,y) < \overline{\psi}n \\ ... \\ 1-\frac{c_j}{2^k} \text{ if } \psi^j \leq \psi_\alpha(x,y) < \psi^{j-1} \\ ... \\ 1/2 \text{ if } \psi_\alpha(x,y) = 0 \\ ... \\ \frac{c_j}{2^k} \text{ if } -\psi^{j-1} \leq \psi_\alpha(x,y) < -\psi^j \\ ... \\ \frac{c_1}{2^k} \text{ if } -\overline{\psi}n < \psi_\alpha(x,y) \leq -\psi^j \\ 0 \text{ if } \psi_\alpha(x,y) \leq -\overline{\psi}n \end{cases}$$

with $\frac{1}{2^k}$ arbitrarily small $k \in N$ and $0 < c_1 < c_2 < ... < c_j$. Our objective is to show that it is the optimal form of the probability assignment function when $\widetilde{F}_n$ is required to belong to the set $\widetilde{\Im}_n$. For that matter, we will show that $F^*_n \in cl(\widetilde{\Im}_n)$ can be uniformly approximated by the functions $F^*_{n\epsilon} \in \widetilde{\Im}_n$ on D $\forall n \in \{1,...,T\}$ to within $\delta > 0$.
To show this, we need to demonstrate that there exists a sequence of functions in $\widetilde{\Im}^*_n$ which converges uniformly on D to F^*.

Definition. We say that a sequence of functions $\{f_n\}$, $n = 1,2,3,...,(\{F^*_{nk}\}$, $k = 1,2,3,...,$in our problem) converges uniformly on X ($D = \left[-\overline{\psi}n, \overline{\psi}n\right]$in our problem) to a function f if for every $\delta > 0$ there is an integer N such that $n \geq N$ ($k \geq N$ in our problem) implies,

$$|f_n(x) - f(x)| \leq \delta$$

for all $x \in E$.

To show that$\{F^*_{nk}\}$converges uniformly to F^*, we apply the Cauchy criterion given in the following theorem,

Theorem. The sequence of functions $\{f_n\}$ defined on E converges uniformly on E if and only if for every $\delta > 0$ there exists an integer N such that $m \geq N$, $n \geq N$, $x \in E$ implies

$$|f_n(x) - f_m(x)| \leq \delta$$

(For a proof, see Rudin "Principles of Mathematical Analysis" page 147.)

Let $\delta > 0$ be given, choose N such that, $\left|F^*_{nk}(\psi_\alpha) - F^*_{nk'}(\psi_\alpha)\right| \leq \delta$ holds. Fix k and let $k' \to \infty$. As $F^*_{nk'} \to F^*$ since $\frac{c_i}{2^{k'}} \to 0$ $\forall i = 1,...,j$as $k' \to \infty$. we have that, $|F^*_{nk}(\psi_\alpha) - F^*(\psi_\alpha)| \leq \delta$ $\forall k \geq N$ and $\psi_\alpha \in E$.

Step 6 Since $F^* \in cl(\widetilde{\Im_n})$ can be uniformly approximated by functions $F^*_{nk} \in \widetilde{\Im}_n$, we can always find a function F^*_{nk} which implements σ^*as an SE by proposition 1 with a $k \in N$ large enough such that, $|F^*_{nk}(\psi_\alpha) - F^*(\psi_\alpha)| \leq \eta$ with $\eta > 0$ arbitrarily small and $n > g(\Psi)$. From step2, we know that the functions $\Delta_\alpha(\widetilde{F}_n)$ are linear in the choice variable for any $\widetilde{F}_n \in \Im_n$. As F^* can be uniformly approximated by $F^*_{nk'} \in \Im^*_n$, we can choose any $\eta > 0$ such that,

$$\left|F^*_{nk'} - F^*\right| < \eta \Rightarrow \left|\Delta_\alpha(\widetilde{F}^*_{nk'}) - \Delta_\alpha(F^*)\right| < \delta,$$

Hence, we conclude that,
$\left|\Delta_\alpha(\widetilde{F}^*_{nk'}) - \Delta_\alpha(F^*)\right| < \delta$ for $\delta > 0$ arbitrarily small and $(\sigma^*_\alpha, \sigma^*_\beta)$ as an SE.

$\forall\delta > 0$ and for any $(\sigma_\alpha, \sigma_\beta) \in \sum\alpha \times \sum\beta$. This is equivalent to,

$$\left|B_\alpha(\widetilde{F}^*_{nk'}) - B_\alpha(F^*)\right| < \delta \ \forall\delta > 0 \text{ and finishes the proof.}$$

Proof of corollary 1

We know from proposition 1 that the optimal solution of the problem under consideration is an extreme point of $cl(\widetilde{\Im_n})$characterized by the *3 steps probability assignment function*. Moreover, if $n \leq g(\Psi)$, it does not exist any bad or good scenarios at any decision nodes. This implies that the optimal probability assignment function $\widetilde{F}^*_n$ need *not* to belong to $\widetilde{\Im}_n$ any longer. Since by definition $\widetilde{\Im}_n \subset cl(\widetilde{\Im}_n)$we can write, $\max\limits_{F_n \in cl(\widetilde{\Im}_n)} \Delta(F_n) \geq \max\limits_{F_n \in \widetilde{\Im}_n} \Delta(F_n) \Leftrightarrow \min\limits_{F_n \in \widetilde{\Im}_n} B(\Psi, n, F_n) \geq \min\limits_{F_n \in cl(\widetilde{\Im}_n)} B(\Psi, n, F_n)$.

Hence, we conclude that $F^*_n \in cl(\widetilde{\Im}_n)$ is the optimal probability assignment function with F^*_n the three step probability assignment function.

A4.

Proof of proposition 3 Let $z_\alpha(\tau) = q_\alpha(\tau) - q_\beta(\tau)$, $\mu_\alpha = \sum\limits_{q \in Q\alpha} p^{e^*_\alpha}_\alpha(q).q$ be the mean production for any period of the game for agent α and $\sum\limits_{q \in Q\alpha} p^{e^*_\alpha}_\alpha (q).[q - \mu_\alpha]$ be the variance of production for agent α within one period. We first look at the probability for the principal to make an error for agent 1 under the sampling mechanism of Dubey et al. The probability for agent 1 to win over a sample of size n while he is *not* the most productive agent over the T periods is,

$$\left(\sum_{\zeta\in C_n} p\left(n\right).\left(1-p\left[\left|\sum_{\tau\in\zeta}^{n} z_1(\tau)-n\left(\mu_1-\mu_2\right)\right|\geq n\left(\mu_1-\mu_2\right)\right]\right)\right).$$
$$\times p\left[\left|\sum_{\tau=1}^{T} z_1(\tau)-n\left(\mu_1-\mu_2\right)\right|\geq n\left(\mu_1-\mu_2\right)\right]$$

Note that every possible sample $\zeta \in C_n$ is chosen with the uniform probability $p\left(n\right) = \frac{1}{C_T^n}$ and, by independence across the time periods, the expression reduces to,

$$\left(1-p\left[\left|\sum_{\tau\in\zeta}^{n} z_1(\tau)-n\left(\mu_1-\mu_2\right)\right|\geq n\left(\mu_1-\mu_2\right)\right]\right)$$
$$\times p\left[\left|\sum_{\tau=1}^{T} z_1(\tau)-n\left(\mu_1-\mu_2\right)\right|\geq n\left(\mu_1-\mu_2\right)\right]$$

Now, under the class of probability assignment functions $F_{Tk}^* \in \Im_T^*$ and *without* any sampling device the probability of this event becomes,

$$\left(\sum_j F_{Tk}^*\left(\psi^j\right) p\left[\sum_{\tau=1}^{T} z_1(\tau)=\psi^j\right]\right).p\left[\left|\sum_{\tau=1}^{T} z_1(\tau)-n\left(\mu_1-\mu_2\right)\right|\geq n\left(\mu_1-\mu_2\right)\right]$$
with $\psi^j < 0, \psi^j \in E$.

By definition of any function $F_{Tk}^* \in \Im_T^*$ and since $\psi^j < 0$, we can write,

$$\sum_j F_{Tk}^*\left(\psi^j\right) p\left[\sum_{\tau=1}^{T} z_1(\tau)=\psi^j\right] = \frac{1}{2^k}\left(\sum_j p\left[\sum_{\tau=1}^{T} z_1(\tau)=\psi^j\right].j\right) \forall j \text{ and }$$
$k \in N$ and $\psi^j \in E$.

Note that this expression is bounded above by $\frac{j.c_j}{k}$. Hence, we have, $\frac{1}{2^k}\left(\sum_j p\left[\sum_{\tau\in\zeta}^{n} z_1(\tau)=\psi^j\right].j\right) < \frac{j.c_j}{2^k}$. which implies that for a k large enough we have, $\frac{1}{2^k}\left(\sum_j p\left[\sum_{\tau\in\zeta}^{n} z_1(\tau)=\psi^j\right].j\right) < \frac{j.c_j}{2^k} < \epsilon$ for $\epsilon > 0$ arbitraly small.

$$1-p\left[\left|\sum_{\tau\in\zeta}^{n} z_1(\tau)-n\left(\mu_1-\mu_2\right)\right|\geq n\left(\mu_1-\mu_2\right)\right]>\epsilon>\tfrac{j.c_j}{2^k}>0.$$

This expression must always hold for $\epsilon>0$ chosen arbitrarily small.

A5.

Proof of proposition 4

Step 1 (Uniform convergence) Using the same argument as in the proof of proposition 3, we know that $F^*\in cl(\widetilde{\Im}_T)$ can be uniformly approximated on $D=\left[-\overline{\psi}T,\overline{\psi}T\right]$ by the functions F^*_{Tk} in $\Im^*_T$.

Step 2 (Continuity) We want to show that the overall probability of winning for each agent $\alpha\in\{1,2\}$ is continuous in the choice variable $F^*_{Tk'}\in\Im^*_T$. From the proof of proposition 3, we know that the functions $p^n_\alpha(\sigma_\alpha,\sigma_\beta,F_n)$ are linear in the choice variable for any $F_n\in cl(\widetilde{\Im_n})$. As every linear function is continuous, we have that, $p^n_\alpha(\sigma_\alpha,\sigma_\beta,F_n)$ is continuous for any $F^*_{Tk'}\in\Im^*_T$ in particular. And, as F^* can be uniformly approximated by $F^*_{Tk'}\in\Im^*_T$, we can choose $\delta>0$ such that,

$$\left|F^*_{Tk'}-F^*\right|<\delta\Rightarrow\left|p^T_\alpha(\sigma_\alpha,\sigma_\beta,F^*_{Tk'})-p^T_\alpha(\sigma_\alpha,\sigma_\beta,F^*)\right|<\epsilon,$$

$\forall\epsilon>0$ and for any $(\sigma_\alpha,\sigma_\beta)\in\sum_\alpha\times\sum_\beta$. This holds in particular for an $\epsilon>0$ small enough.

Step 3 We show that *without* any sampling device, the overall number of periods T maximizes the probability of winning. By step 2 above, we have just to show that for the profile $\sigma^*$the following inequality holds, $p^T_1(\sigma^*_\alpha,\sigma^*_\beta,F^*)>p^n_1(\sigma^*_\alpha,\sigma^*_\beta,F^*)$ for any $1\leq n<T$. To this end, we apply the following inequality due to Bernstein,

Let $\mu_\alpha=\sum_{q\in Q\alpha}p_\alpha^{e^*_\alpha}(q).q$ be the mean production for any time period of the game for agent α and $v_\alpha=\sum_{q\in Q\alpha}p_\alpha^{e^*_\alpha}(q).\left[q-\mu_\alpha\right]$ be the variance of production for agent α within one period. Let $z_1(\tau)=q_1(\tau)-q_2(\tau)$ be independent random variables such that,

$p\left[\left|\sum_{\tau\in\zeta}^{n} z_1(\tau)-n\left(\mu_1-\mu_2\right)\right|\leq\gamma\right]=1$ $\tau=1,...,n$ with $\gamma<\infty$. and $k=\mu_1-\mu_2>0$. Then, for agent 2, the probability of winning under a sample of size n is at most,

$$p\left[\left|\sum_{\tau\in\zeta}^{n} z_1(\tau) - n(\mu_1 - \mu_2)\right| \geq n(\mu_1 - \mu_2)\right] \leq 2\exp(\frac{-nk^2}{2(v_1+v_2)+2/3kn}).$$

Hence at least, agent 1 gets a probability of $1 - 2\exp(\frac{-nk^2}{2(v_1+v_2)+2/3kn})$. Thus, we get,

$p_1^T(\sigma\alpha^*, \sigma\beta^*, F^*) > p_1^n(\sigma\alpha^*, \sigma\beta^*, F^*)$ for $1 \leq n < T$.

References

[1] Che, Y-K., Gale, I., 2000. "Difference-form contests and the robustness of all pay auctions" 30, 22-43 *Games and Economic Behavior.*

[2] Dubey, P., Haimanko, O.," Optimal scrutiny in multi-period promotion tournament," Forthcoming. *Games and Economic Behavior.*

[3] Dubey, P., Haimanko, O., 2001."Envy and the optimality of tournament," DP#250 of center for rationality and Interactive Decision Theory. Hebrew University, Jerusalem.

[4] Dubey, P., Wu, C-W., 2001." Competitive prizes: when less scrutiny induces more effort," *Journal of Mathematical Economics* 36, 311-336.

[5] Green, J.R., Stokey, N.L., 1983." A comparison of tournaments and contracts,". *Journal of Political Economy* 91, 349-364.

[6] Grossman, S.J., Hart, O., 1983. "An analysis of the principal agent problem,". *Econometrica* 51, 7-45.

[7] Hvide, H.K., Kristiansen, E.G., " Risk taking in selection contests," Forthcoming *Games and Economic Behavior.*

[8] Lazear, E.P., Rosen, S., 1981."Rank-order tournament as optimum labor contracts,". *Journal of Political Economy* 89, 841-864.

[9] Mookherjee, D., 1984. "Optimal incentive schemes with many agents," *Review of Economic Studies* LI, 433-446.

[10] Munkres, I.,1975 "Topology: A First Course". Prentice Hall.

[11] Rudin, W.,1976."Principles of Mathematical Analysis," Third edition. Mac Graw-Hill,Inc.