UNIVERSIDAD CARLOS III DE MADRID working papers

Working Paper 04-07
Business Economics Series 05
February 2004

Departamento de Economía de la Empresa
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34-91) 6249608

MORAL HAZARD IN TEAMS WITH LIMITED PUNISHMENTS AND MULTIPLE OUTPUTS

Jose E. Quintero Jaramillo*[1]

Abstract

This paper studies incentive provision with limited punishments. It revisits the moral hazard problem with risk neutral parties and solves for optimal compensation schemes in situations where agents' participation is implied by a limited liability constraint. Providing minimum cost incentives to teams or individuals requires awarding high bonuses only when extreme performances are observed. Even when the first-best is attainable, the principal may prefer to induce more (or less) effort than it is socially desirable because she only cares about the marginal cost of motivation. With positive production externalities joint bonuses are optimal. With limited liability on the principal's side, the optimal scheme becomes a tournament---even in the absence of externalities. The paper also looks at conditions that favor one incentive scheme over another when agents adapt their strategies as information becomes available.

Keywords: Moral Hazard, Limited liability, Tournaments, Compensations Schemes.

JEL classification: D82, J33

*I am grateful to Ilya Segal and Jon Levin for advice. I thank Simon Board, David Luenberger, Hanno Lustig, Paul Milgrom, Luis Rayo, Robert Wilson, and seminar participants at Stanford for valuable discussions and comments.

[1]Department of Business Administration. Universidad Carlos III de Madrid. C/Madrid, 126. 28903 Getafe (Madrid, Spain). E-mail: jose.quintero@uc3m.es

Moral Hazard in Teams With Limited Punishments and Multiple Outputs*

Jose E Quintero†
Economía de Empresa, Universidad Carlos III de Madrid.

February 3, 2004

Abstract

This paper studies incentive provision with limited punishments. It revisits the moral hazard problem with risk neutral parties and solves for optimal compensation schemes in situations where agents' participation is implied by a limited liability constraint. Providing minimum cost incentives to teams or individuals requires awarding high bonuses only when extreme performances are observed. Even when the first-best is attainable, the principal may prefer to induce more (or less) effort than it is sociably desirable because she only cares about the marginal cost of motivation. With positive production externalities joint bonuses are optimal. With limited liability on the principal's side, the optimal scheme becomes a tournament—even in the absence of externalities. The paper also looks at conditions that favor one incentive scheme over another when agents adapt their strategies as information becomes available.

*I am grateful to Ilya Segal and Jonathan levin for advice. I would also like to thank Susan Athey, Luis Rayo, Shikar Ranjan, and seminar participants at Stanford for valuable discussions and comments.

†C/ Madrid 126, Getafe-Madrid 28903, Spain. EMAIL: jose.quintero@uc3m.es.

1 Introduction

Over the past thirty years there has been a great effort to understand how firms provide incentives to their employees. Although there is strong evidence regarding the importance of economic incentives in motivating workers, there is little agreement on whether firms use the optimal schemes predicted by the theory (see Prendergast (1999) and references therein). First, optimal trade-off between insurance and incentives (Holmstrom (1979, 1982), Shavell (1979)) results in more complicated contracts than the ones observed in practice. Second, specific models that yield simpler contracts like piece rates (Holmstrom and Milgrom (1987)), stock options (Innes (1990)), efficiency wages (Shapiro and Stiglitz (1984)), or promotion tournaments (Lazear and Rosen (1981), and Green and Stokey (1983)) do not offer a systematic explanation for why an employer should select one contract over another.

This paper sheds some new light on the conditions affecting the choice of compensation contracts. The point of departure is the non binding participation constraints of a risk neutral agent. The optimal contract solution is stark, offering a very high bonus for extreme (but rare) performances. Adding different institutional constraints (such as limited liability for the principal and no arbitrage) yields more commonly observed contracts such as bonuses and stock options. Through extending the model to multiple agents and allowing for externalities in agents' production, it is shown that individual and joint bonus schemes, as well as tournaments arise as minimum cost incentive schemes.

In the standard principal-agent analysis, contract design is driven by (i) a trade-off between risk-sharing and incentive provision, and (ii) matching the reservation utility of the agent. In the standard model, the cost of providing incentives depends largely on (ii), the agent's participation constraint. I drop risk aversion and individual rationality considerations, replacing them with a limited liability condition that bounds payments to the agent from below. I examine the cost of putting in place an incentive scheme that achieves a given effort target.

But why is the problem relevant for practice in spite of the above simplifications? First, because it is reasonable to assume that firms try to minimize cost when designing compensation incentives. Second, in a limited liability world agents are not exposed to dramatic punishments which relaxes their need for insurance. Third, if motivation for different tasks can be viewed as being additive, the cost of giving incentives for an additional task may be examined independently. If this is the case, marginal incentives for a secondary task would not necessarily need to be subject to participation restrictions. Moreover, if incentives come from many dimensions of work, one would expect risk neutrality to be a good approximation of agents behavior with respect to income changes due to one of the tasks. For example, a firm may use efficiency wages to maintain a hard working team. By offering higher wages than the competition, the firm induces effort from agents that fear being let go. However, suppose the firm desires to give additional incentives, for example, to workers that bring new sales. Although insurance and participation considerations may be important for the efficiency wage scheme, they play no role for designing additional incentives. Hence, the latter should respond to a minimum cost criterion.

Section 2 deals with the case of a principal contracting with a single agent. Mirrlees (1974, 1999) conjectures that penalizing individuals for rare events that are strong indicators about deviations is a powerful incentive. Holmstrom (1982) formalizes this idea showing that, under restrictive conditions, the first-best effort can be induced at negligible cost using either penalties or bonuses. The first result is an extension of Holmstrom and Mirrlees' observation. Here I relax the conditions (extreme events need not be strong evidence of the effort choice) and show that such contracts may still minimize the cost of motivating agents, even though the cost may not be negligible. The contract pays a high bonus only when the best possible outcome is observed. This is a zero probability event in the model, thus the bonus has to be infinite; however, its expected value exists and the minimum cost of inducing any action is associated with a limit.

Given the unpractical nature of optimal contracts, the paper goes on to consider restrictions on the contract space corresponding to natural institutional constraints. In particular, with double-sided limited liability (compensation schemes bounded above and below), the optimal contract becomes a more reasonable form of bonus, i.e. a one step function. Imposing self-financing and no arbitrage conditions yields similar findings to those of Innes (1990), and Oyer (2000). Specifically, corollary 7 shows that debt—equivalently, stock option plans—are optimal incentives when both parties require their payments to be non-decreasing in output.

Section 3 extends the model by considering multiple agents. In the same vein that Mookherjee (1984), the paper allows for externalities among agents, and for individual output signals. The difference in results is mostly due to the emphasis here on limited liability considerations, risk-neutrality and, to a lesser extent, on using a continuous model (in efforts and outputs). The same result as in the single agent case obtains: large bonuses for rare events minimize the cost of providing incentives. Proposition 11 shows that under positive externalities the most effective compensation scheme is a joint bonus; while if externalities are negative it is an extreme form of relative performance evaluation (RPE)—a tournament that requires a large gap between the winner and the runner-up.

More interestingly, Proposition 16 shows that adding a limited liability condition to the principal, tournaments with a single prize are the unique optimal scheme in all cases. This result parallels the findings of Levin (2002) in a dynamic setting, and is new to the theory of tournaments. Even in the case of independent outputs, as well as with complementarities, the principal benefits from using all other agents' outputs in each agent's compensation scheme. This result seems to defy the logic of the sufficient statistics results of Holmstrom (1979, 1982). The explanation for this discrepancy is quite simple: even though output functions are independent, the principal's limited liability condition affects all compensation schemes simultaneously. Ex-ante, the most efficient way to induce effort is to reward very high outputs. Hence, it is in the best interest of the principal to use the ex-post information, and allocate the scarce motivational resources by comparing outputs.

The literature on tournaments does not present a strong argument for why such compensation schemes are used so broadly. The pioneering work is Lazear and Rosen (1981); it shows that contests can achieve the same results as other incentive schemes, and argue that implementing contests requires less monitoring and information than other schemes. Probably the strongest and most common argument in favor of RPE schemes in general is that they filter noise that affects all agents in the same way (see Green and Stokey (1983), and Nalebuff and Stiglitz (1983a)). However, Mookherjee (1984) shows that when agents are risk-averse, rank order tournaments are optimal schemes *only* under the restrictive condition that rank order is a sufficient statistic. This paper then adds an argument for tournaments based on the cost of implementing incentive schemes. Moreover, this seems to answer what Prendergast (1999 p.36) finds puzzling about why tournaments are much more prevalent than the theory would predict.

Before concluding in Section 5, Section 4 deals with the agents' reaction to partial revelation of information. There has been little work devoted to understanding the adaptation of behavior halfway through the monitoring period. A number of empirical studies show evidence of agents gaming the incentive schemes presented to them (e.g. Oyer (1998), Asch (1990), Ehrenberg and Bognanno (1990), Zenger (1992)). Holmstrom and Milgrom (1987) makes an important theoretical contribution on this issue showing that piece-rate schemes—which agents find hard to game—are optimal in certain circumstances. Cabral (2002) and Oyer (2000), among others, present theoretical examples of such gaming showing that under non-linear incentive schemes, under-performers would tend to choose risky actions. Similar ideas can be found in the literature on R&D races (e.g. Athey and Schmutzler (2001)).

I use a two-agent version of the model to explore agents' reaction to partial interim revelation of information. I investigate agents' behavior under three compensation schemes suggested in Sections 2 and 3: individual performance evaluation (IPE), joint performance evaluation (JPE), and a particular form of relative performance evaluations (RPE): tournaments. On the negative side, the results confirm the broad intuition that such schemes are not robust and that agents will game the system. On the positive side, the analysis reveals some conditions that favor one scheme over another. Specifically, positive correlation in production favors a tournament, while negative correlation favors a joint bonus scheme. In a tournament, agents exert high levels of effort when their competitors are close. If there is a large gap, the laggard finds it too hard to catch up, the leader can comfortably stay ahead, and none of them will exert much effort. Similarly, under a bonus scheme high levels of effort will be observed when there is a reasonable chance to get the reward. If a worker has already reached the reward level she may stop working; if she is too far behind, she will see no point in trying hard. If output is positively correlated, there is a better chance that competition will be tight at the time information is revealed. If the correlation is negative, there is a good chance that a bonus for joint output will keep incentives high.

2 Motivating a Single Agent

A firm's output x depends (at least partially) on an agent's choice of action a. The firm cannot observe the agent's action, nor can it infer the action perfectly from the observed output. More specifically, $X(a)$ is a random variable with support in $[0, \bar{x}]$ with continuous differentiable density function $f(x|a) > 0$.

The action is assumed to be a scalar $a \in A \subset \mathbb{R}$. Moreover assume that a higher action yields a better distribution of output in the likelihood ratio order sense: $a > a'$ implies $\frac{f(x|a)}{f(x|a')}$ is a non-decreasing function of x; and write it as: $f(\cdot|a) \succeq_{LR} f(\cdot|a')$. The choice of action is costly; $c(a)$ represents the disutility to the agent of choosing action a, throughout it is assumed that $c \geq 0; c'(0) = 0,\ c'' > 0$, and if $\bar{a} = \sup\{a \in A\}$, then $\lim_{a \uparrow \bar{a}} c(a) = \infty$.

Consider the following program:

$$\max_{a,s(x)} \int_0^{\bar{x}} (x - s(x)) f(x|a)\, dx \tag{1}$$
$$\text{s.t. (i)} \int_0^{\bar{x}} s(x) f(x|a)\, dx - c(a) \geq \bar{u}$$
$$\text{(ii) } a \text{ solves } \max_{\hat{a}\in A} \int_0^{\bar{x}} s(x) f(x|\hat{a})\, dx - c(\hat{a})$$
$$\text{(iii) } s(x) \geq L$$

This is a typical moral hazard problem between risk neutral principal and agent. The principal seeks an output contingent contract $s(x)$ to motivate the agent. The agent is free to choose the action a from the set $A \subset \mathbb{R}$. Disregarding the limited liability constraint (iii) it is well known that a simple solution is to sell the project x to the agent at the maximum price she would be willing to pay. That is $s(x) = x - k$, where k solves (i) with equality. Of course, there are other solutions that induce the same effort level, while giving the agent her outside option. With risk-averse agent (and principal), the issue becomes one of trading off risk and incentives. Similarly, when the limited liability constraint (iii) is taken into account, the design of the contract becomes crucial; here it is a question of using incentive instruments in the most efficient way. This effect is most striking when the limited liability constraint implies the participation constraint (i) –e.g. if $L \geq \bar{u} + \max c(a)$.

An useful way of solving moral hazard problems, following Grossman and Hart (1983), is to investigate first the optimal contract that induces a particular level of effort. Then one considers what effort level will be desirable to induce. This intuitive approach fits perfectly here since it will emphasize the shape of a contract that achieves a particular effort goal at minimum cost.

A further simplification of the problem is obtained using the so-called first-order approach, which replaces the incentive compatibility constraint (ii) with its first order condition. Once a solution to the problem is found, one can verify that it satisfies the global incentive compatibility constraint of the agent.

I start by considering the following version of the problem,

$$\min_{s(x)} \int_0^{\bar{x}} s(x) f(x|a)\, dx \tag{P}$$
$$\text{s.t.} \quad \text{(i)} \quad \int_0^{\bar{x}} s(x) f(x|a)\, dx - c(a) \geq \bar{u}$$
$$\text{(ii)'} \quad \int_0^{\bar{x}} s(x) f_a(x|a)\, dx = c'(a)$$
$$\text{(iii)} \quad s(x) \geq L$$

Note that the objective was changed, so that the new program minimizes the payment to the agent, and the only decision variable is the payment schedule (effort has been fixed). Constraint (ii)' is the first order condition for (ii). In fact, most of the results below deal with a relaxed version of the moral hazard program, one that neglects the participation constraint.

$$\min_{s(x)} \int_0^{\bar{x}} s(x) f(x|a)\, dx \tag{RP}$$
$$\text{s.t.} \quad \text{(ii)'} \quad \int_0^{\bar{x}} s(x) f_a(x|a)\, dx = c'(a)$$
$$\text{(iii)} \quad s(x) \geq L$$

There are two reasons for disregarding participation considerations. First, working with the less constrained problem will highlight the features of contract design that deal exclusively with incentive provision. In other words, the shape of the optimal contract will represent the least costly way of providing incentives, rather than the optimal way of motivating the agent while ensuring a minimum transfer of utility.

Second, in this way the model captures some applications that have not been addressed in the literature. There are many circumstances in which it is important to motivate agents whose participation in the contract is somehow guaranteed by the limited liability condition. Take the case of locked-in employees. A firm may have some workers that have a very specific experience, whose outside option is much lower than their current wage, but are highly valuable for the firm. Athletes and other potential superstars face a similar dilemma at the beginning of their careers. Their outside alternatives may be represented by similar jobs to those they perform anyhow while training to make it into professional leagues.

As will be seen, with a binding participation constraint there will be a larger class of solutions. That is, optimal contract-design becomes more flexible when the agent's participation is not guaranteed by the limited liability constraint. The following proposition shows that whenever the participation constraint binds, program P has a one step (bonus contract) solution.

Proposition 1 *Given effort a, if*

$$\bar{u} + c(a) > L + c'(a) \frac{f(\bar{x}|a)}{f_a(\bar{x}|a)} \tag{2}$$

then, there exists some $\tilde{x} \in [0, \bar{x}]$, $k \geq 0$, and $M \geq L$ such that problem P has a solution of the form[1]

$$s(x) = M + k 1_{x \geq \tilde{x}} \tag{3}$$

Proof. in the appendix. ■

Of course, the solution proposed above is not unique. For example, it is well known (Harris and Raviv (1976)) that when L is sufficiently low, selling the project to an agent achieves the efficient level of effort. Similarly, different effort targets can be achieved by selling a share of the project when the agent's liability is flexible enough (L is small). The contracts proposed here are similar to those proposed by Mirrlees (1974) and Holmstrom (1982), where they make the point that first best can be achieved at negligible cost under more restrictive conditions. In this direction, proposition 1 suggests a lower bound for the cost of motivating agents in a more general framework; proposition 3 will show that this bound is tight.

Notice that in the standard moral hazard problem considered in the literature –e.g. Holmstrom (1979, and 1982)– the participation constraint determines the expected transfer the agent receives. The last proposition shows that there is no loss to restrict attention to bonus (one step) contracts to achieve optimality in such cases. When the principal does not need to induce participation, the relevant question becomes at what price can she implement a desired level of effort. Such questions are particularly important in the context of motivating a team (which will be handled in section 4); but most of the intuition will come from the single agent case.

Proposition 3 shows that when the IR constraint does not bind, the least cost of motivating the agent is achieved by an extreme form of bonus contract: one that would only pay for the best possible outcome. Under the assumptions so far, such an event has zero probability and therefore the bonus must be infinite. In reality such a contract is not feasible. However, the proposition will show that the bound proposed in proposition 1 is tight; validating the simple one-step approximations used in proposition 1. Moreover, it gives rise

[1] 1_A will denote the indicator function of the set A. That is $1_A(x) = 1$ if A occurs, and 0 otherwise.

to the structure of optimal contracts under several institutional constraints as the corollaries will illustrate.

Before proceeding, it will be useful to recall the definition of the Dirac-delta function –taken from Luenberger (1964 p. xvi). This is the mathematical representation of the stark solution in proposition 3.

Definition 2 *The Dirac delta function* $\delta(\cdot)$ *is defined by the relation*

$$\int_a^b f(t)\,\delta(t)\,dt = f(0)$$

for every continuous function f *in* $[a,b]$, *where* $0 \in (a,b)$.

Proposition 3 $s(x) = L + \frac{c'(a)}{f_a(\bar{x}|a)}\delta(x-\bar{x})$ *is a solution to problem (RP). That is, a Dirac function that recognizes only a "super-bonus" for the best possible outcome is the least expensive way of providing high-powered incentives. Moreover, if* $f(\bar{x},a)$ *is a concave function of* a, *then the global incentive constraint (ii) is also satisfied.*

Proof. in the appendix. ■

The proof is very intuitive. The first-order-condition of the incentive constraint determines the *quantity* of incentives required. Incentives need to take the form of a premium wage for some observed output levels. There is a shadow cost of providing incentives at each level of performance, and the cost is lowest when observed output is a better signal of more effort. The likelihood ratio ordering assumption implies that the best signal for increased effort are the very best performances. Hence, offering all the required incentives for the best possible result is the least costly way to induce the agent to exert up to the desired level of effort.

As corollary, the minimum (expected) cost of motivating the agent to exert effort a is

$$m(a) := c'(a)\,\frac{f(\bar{x}|a)}{f_a(\bar{x}|a)}.$$

An interesting point regarding the type of super-bonus contracts of proposition 3 is that the principal may implement an inefficient level of effort (either above or below the first best). Specifically, given the minimum cost of implementing effort a, the principal will implement a^* that solves

$$\max_{a\in A} E[x|a] - m(a).$$

Whereas a social planner would implement the first-best level of effort: a^{FB}, a solution to

$$\max_{a \in A} E[x|a] - c(a).$$

Consequently, the principal will induce the agent to work too much with respect to the social optimum $\left(a^* > a^{FB}\right)$ when $m'(\cdot) < c'(\cdot)$. And she will induce less effort than desirable if $m'(\cdot) > c'(\cdot)$; both cases being plausible.

This could be the case of some sports, like cycling, where being one step ahead of anti-doping policies is the name of the game. The rules of competition (explicit and implicit) force athletes to go beyond their capabilities; the additional rents from more spectacular performances outweigh the additional cost of *inducing* more effort than desirable. That is, even though social surplus decreases for a above the efficient level; the rents that the agent claims, $m(a) - c(a)$, may decrease faster with respect to induced effort a; consequently, the principal finds in her own advantage to induce more effort than desirable. The following corollary shows, with a mathematical example, that such cases are in fact possible.

Corollary 4 *The level of effort that the principal induces, a^*, may be lower or **higher** than the first-best level a^{FB}.*
There are rare cases in which the principal desires to induce more effort than first-best; in the following example I choose appropriate functional forms of costs and distributions to show that over production of effort may arise. Let $c(a) = \frac{e^a - 1}{3}$, and $f(x, a) =$

With a binding participation constraint, there would still be a solution of the form found in proposition 3—although the minimum payment may need to be higher than L. The following corollary relates Propositions 1 and 3 by showing that one-step contracts of the type used in Proposition 1 are constrained optima when limited liability affects both parties.

Corollary 5 *Suppose the transfer from the principal cannot exceed some amount $W > L$. Then, if there exists a solution to RP with the additional constraint, it has the form*

$$s(x) = \begin{cases} L & x < x^* \\ W & x \geq x^* \end{cases}$$

for some $x^* \in [0, \bar{x}]$.

Proof. (sketch) With the additional constraint, the program becomes:

$$\min_{s(x)} \int_{\hat{x}}^{\bar{x}} (s(x) - L) f(x|a)\, dx$$
$$\text{s.t. (ii)'} \quad \int_{\hat{x}}^{\bar{x}} (s(x) - L) f_a(x|a)\, dx = c'(a)$$
$$\text{(iii)' } W - L \geq s(x) - L \geq 0$$

As before, the solution consists of using the region where $\frac{f(x|a)}{f_a(x|a)}$ is minimum up to the point that (ii)' is satisfied. Hence x^* will be defined by

$$\int_{x^*}^{\bar{x}} (W - L) f_a(x|a)\, dx = c'(a)$$

■

Under many institutions, contracts take the form of simple sharing rules. For example in sharecropping, parties decide how to split the value of the crops, but no party makes payments greater than the value of the harvest. The same holds for debt contracting. In light of the above results, it is not surprising that the least costly way to motivate the agent is to give her all the output in good states of the world (high values of x), and zero in bad states. Innes 90 obtains such a result, which can be stated as a corollary to proposition 3.

Corollary 6 *(Innes 90) Consider program RP with the additional constraint*

$$\textit{(iii)''} \quad 0 \leq s(x) \leq x.$$

The solution has the form

$$s(x) = \begin{cases} 0 & x < x^* \\ x & x \geq x^* \end{cases}$$

for some x^*.

These type of contracts are hardly seen in reality. Both parties would have a strong incentive to misrepresent information when output is close to the threshold level x^*. A reasonable condition that one may want to impose is that the share of both parties should be non-decreasing in x. This could be interpreted as a no-arbitrage condition: neither party benefits from throwing away output, or from buying additional amounts. The new requirement is that both $s(x)$, and $x - s(x)$ must be non decreasing. Innes 90 [13] also studies this case and finds that the optimal incentive scheme is a debt contract; that is $s(x) = \max\{0, x - d\}$ for some d. As with the previous corollary, the proof is a simple extension of proposition 3 and will be omitted.

Corollary 7 *(Innes 90) Consider program RP with the additional constraints*

$$(iii)''' \quad 0 \leq s(x) - s(x') \leq x - x' \text{ for all } x > x'$$

The solution is a debt contract $s(x) = \max\{0, x - d\}$ *for some* d.

Obviously, under such constraints it may be impossible to implement high levels of effort. In particular, a debt contract with a positive repayment value ($d > 0$) cannot achieve the first-best level of effort. This is in contrast to the type of contracts described in corollary 6, which can implement first-best (and even higher) levels of effort.

3 Motivation for Teams

This section extends the model to incentive schemes for a team. I shall consider a principal that observes individual output signals from a group of agents and who is interested in the sum of their outputs. The model will allow for positive as well as negative externalities among agents' production technologies. Some of the results also apply to situations where the number of signals observed differs from the number of agents. For example, where there are different dimensions to each agent's output and the principal is interested in the sum of such components. However, as in the previous section, the effort choice is just a scalar.

Start by considering a simple extension of the previous model, where n agents require a costly action to produce some output. Each agent $i \in \{1, 2, ..., n\}$ chooses her effort a_i. I use the notation $a = (a_1, ..., a_n) = (a_i, a_{-i})$.The cost of effort for agent i is given by $c_i(a_i)$. Agent i's output, $x_i \in [0, \bar{x}_i]$, is distributed according to the c.d.f. $F^i(\cdot|a_i, a_{-i})$, with differentiable density $f^i(\cdot|a_i, a_{-i})$, which is assumed positive on $[0, \bar{x}_i]$.

Each agent's effort is productive, and outputs are independently distributed given effort a. That is, all the information that x_j can have about x_i is contained in the effort vector a. Formally, let $X = \sum_i x_i$ and $h(\cdot|a)$ be the density of X given a, with corresponding c.d.f. $H(\cdot|a)$.

Condition 8 (Productivity)
a) For every i, $a_i > a_i'$ *implies that* $f^i(\cdot|a_i, a_{-i}) \succeq_{LR} f^i(\cdot|a_i', a_{-i})$
b) For every i, $a_i > a_i'$ *implies that* $h(\cdot|a_i, a_{-i}) \succeq_{LR} h(\cdot|a_i', a_{-i})$

Condition 9 (Independence) *Given effort* a, $\{x_1, ..., x_n\}$ *are independent random variables. That is the joint density of* $\{x_1, ..., x_n\}$ *given effort* a *is:*

$$\mu(x|a) = \prod_i f^i(x_i|a)$$

The productivity condition says that individual and total output have better distributions (in the likelihood ratio order sense) when any one agent increases her effort, while the rest of the team works at a fix level of efforts. Note that part b) implies that, regardless of the externalities that i imposes on others' signals, the input of agent i is overall productive. The independence condition implies that there is no common shock that affects agents' productivity. Externalities are represented through the effect that an agent's effort has over other agents' output. More precisely, I will use the following definitions.

Definition 10 *Agent i has positive (negative)* externalities *over k if $a_i > a_i'$ implies that $f^k(\cdot|a_i, a_{-i}) \succeq (\preceq)_{LR} f^k(\cdot|a_i', a_{-i})$. A* team *of agents has positive (negative)* externalities *if every agent has positive (negative) externalities over every other agent in the team.*

The principal is interested in the sum of outputs: $X = \sum_i x_i$. As before, I shall look into the question of how can the principal induce a target level of effort $a^* = (a_1^*, ..., a_n^*)$ at minimum expected cost. When the form of the incentive schedule is not constrained, the solution has the same extreme attributes as in the single agent case: each agent would be rewarded only when the best possible signal is observed.

Given that output from all agents are informative about x_i, in general, the reward of agent i will depend on all signals. Let $s_i(x)$ be the payment to agent i when x is observed. Each agent chooses effort to maximize expected reward, while assuming other agents exert effort a_{-i}^*. Agent i solves the problem:

$$\max_{a_i} E\left[s_i(x)|a_i, a_{-i}^*\right] - c_i(a_i).$$

The first order condition can be written as:

$$\begin{aligned} \int s_i(x)\,\mu_{a_i}\left(x|a_i, a_{-i}^*\right) dx_1...dx_n &= c_i'(a_i), \quad \text{or} \qquad (4) \\ \int s_i(x)\,\mu\left(x|a_i, a_{-i}^*\right)\left[\sum_j \frac{f_{a_i}^j\left(x_j|a_i, a_{-i}^*\right)}{f^j\left(x_j|a_i, a_{-i}^*\right)}\right] dx_1...dx_n &= c_i'(a_i) \end{aligned}$$

The bracketed expression is a measure of how informative is the observed signal about agent i's effort choice. It will reappear often enough; for convenience I will label it $\lambda_i(x, a)$; that is:

$$\lambda_i(x, a) := \sum_k \frac{f_{a_i}^k(x_k|a)}{f^k(x_k|a)}.$$

As in proposition 3, the optimal incentive scheme is going to rely on paying the agent extremely well for realizations x that yield high values of λ_i. Assuming condition 4 is sufficient for each agent's problem; the principal's program can be written as follows.

$$\min_{\{s_i\}_i} \sum_i E\left[s_i(x)\,|a^*\right] \tag{5}$$
$$\text{s.t. } \forall i \ (4)$$
$$\forall i \ \ s_i(x) \geq L_i \tag{6}$$

The participation constraints have been omitted. I restrict attention to the case in which participation is implied by the limited liability constraints (6). As in the one agent case, binding participation constraints just give the principal more flexibility in designing the optimal incentive scheme. The following result generalizes proposition 3 to a multi-agent setting; the proof is very similar.

Proposition 11 *Under the productivity and independence assumptions, the solution to problem 5 is to offer each agent i a "super-bonus" when* $x \in \arg\max_x \lambda_i(x, a^*)$.
a) If agent i's choice of effort does not affect signal x_j*, then i's optimal compensation can be made independent of j's signal.*
b) Under team positive externalities, a solution is to pay all agents the super-bonus only when $(\bar{x}_1, ..., \bar{x}_n)$ *is observed.*
c) Under team negative externalities, a solution is to pay agent i the super-bonus only when $(\bar{x}_i, 0)$ *is observed.*

Proof. see appendix. ■

It is worth noting that if the principal is bound to use only i's signal to compensate agent i, proposition 3 still holds: the principal would give a large bonus for the very best realizations of the individual signal. Such schemes are often called individual performance evaluation (IPE), whereas the schemes that use all agents' information are called joint performance evaluation (JPE). Clearly the cost of attaining a given level of efforts is smaller under JPE, than under IPE. In particular, from proposition 3 , the minimum cost of inducing effort a^* through an IPE schemes is:

$$\sum_i \frac{c_i'(a_i^*)}{\frac{f^i_{a_i}(\bar{x}_i|a^*)}{f^i(\bar{x}_i|a^*)}} \tag{7}$$

Whereas with a JPE scheme, the minimum cost follows from proposition 11 and is

$$\sum_i \frac{c_i'(a_i^*)}{\sum_j \frac{f_{a_i}^j(\tilde{x}_j^i|a^*)}{f^j(\tilde{x}_j^i|a^*)}} \tag{8}$$

with $\tilde{x}^i \in \arg\max_x \lambda_i(x, a^*)$ for each i.

Given that $\frac{f_{a_i}^j(\tilde{x}_j^i|a^*)}{f^j(\tilde{x}_j^i|a^*)}$ is non-negative at the highest point, it is clear that the JPE schedule will dominate individual schedules. In the presence of any kind of externalities –i.e. when at least one such ratio is not always zero for $j \neq i$– the motivation cost will be strictly lower under JPE.

As it could be expected, for teams with positive externalities, the result suggests a large joint bonus when observing excellent performances. Under positive externalities (complements), this translates into joint bonuses; under negative externalities (substitutes), it translates into an extreme form of tournament. A tournament that requires the winner to have a big advantage over the runner-up to claim the prize.

Although joint bonuses are not uncommon in modern organizations, they are definitely not the norm either (see Lazear (1995), Baron and Kreps (1999)). The suggested solution for the case of negative externalities would be hard to implement: Companies do not deny a bonus to a worker just because her peers did not perform poorly! Workers interacting under such conditions would have to spend a great deal of time watching their backs, making sure other employees are not sabotaging their work. Instead, individual bonuses and promotion tournaments seem to be widely used in practice. It is worth then to explore under what conditions the cost of an IPE scheme (individual bonuses) will outperform regular tournament schemes. The following definitions introduce the type of tournaments that will be considered, and a notion of symmetry among agents.

Definition 12 *i) A* tournament *among* n *agents is an incentive scheme that ranks agents' outputs from highest to lowest and pays a prize to each that depends solely on the agent's ranking.*
ii) A tournament with a standard m, *is a tournament that only pays prizes to agents that produce above* m.

A tournament then consists of a set of prizes $\{B_1, ...B_n\}$, and payment functions defined by:
$s_i(x) = B_k$ if and only if x_i is the k-th order statistic from $\{x_{1,}..., x_n\}$ for all i and k.

Definition 13 *Given $a \in A^n$, and $i \neq j$; let a^{ij} be the vector of efforts that results from interchanging the i-th and j-th components*[2] *of a. A team of agents is said to be* symmetric *if*
i) for all i, $c_i(\cdot) = c(\cdot)$, and $\bar{x}_i = \bar{x}$,
ii) for all k and every $i, j \neq k$; $f^k(x_k|a) = f^k(x_k|a^{ij})$, and
iii) for all $z \in [0, \bar{x}]$, and all $i, j \in \{1, ..., n\}$; $f^i(z|a) = f^j(z|a^{ij})$

A symmetric team requires that all agents have the same production functions, and that they cause identical externalities to all other agents.
The following results will be considered only in a simplified version of the model with two symmetric agents, and tournaments with a single prize[3]. Nalebuff and Stiglitz (1983) and Moldovanu and Sela (2001) provide arguments to have a small number of prizes in a tournament. Generally one prize is sufficient, and I will only consider this case.

Proposition 14 *With two symmetric agents with positive externalities, the cost of inducing efforts $a^* = (a_1^*, a_2^*)$, $a_1^* = a_2^*$, via a single-prize tournament with a standard is higher than in the case of using individual bonuses. With negative externalities, the tournament can outperform individual bonuses.*

Proof. See appendix. ■

The proof is straight forward for the case of positive externalities. Given the standard m, one computes the cost of implementing a^* by means of a tournament and compare it to (7). The case of negative externalities requires an example which is provided in the appendix. The intuition is simple: with sufficiently large negative externalities, a tournament approximates better the full (JPE) optimum.

Most of the literature on tournaments emphasizes the benefits of contests, and of relative-performance-evaluation (RPE) schemes in general, because of two reasons. The first is that tournaments filter out noise that affects all players (Green and Stokey (1983), and Nalebuff and Stiglitz (1983)). If, for example, weather conditions have a significant effect on the productivity of a team of workers; making payments depend only on individual outputs may not be as effective an incentive as relying on comparative measures. Workers may realize they will not be fairly compensated when facing adverse conditions; and therefore decide to work only when the weather is favorable. This effect will not be considered in depth here.

[2] That is, $a_j^{ij} = a_i$, $a_i^{ij} = a_j$, and $a_k = a_k^{ij}$ for $k \notin \{i, j\}$.

[3] For regular tournaments with n agents it always suffice to consider just $n-1$ prizes. However, in tournaments with a standard having as many prizes as agents makes a difference.

The second argument for RPE is based on substitution effects. If a job requires to accomplish several tasks that all agents can handle, it makes sense to have an incentive system based on comparing their relative achievements. The best example is probably the case of sports competitions that are organized in the form of a tournament, and give prizes according to the final ranking of competitors. The negative externalities considered in the model capture such effects.

In addition to these two arguments; supporters of RPE schemes like to argue that informational requirements are smaller than in other incentive schemes. The central point being that it requires less monitoring to obtain a comparative measure of performance than to obtain an absolute one. Supporting this type of heuristic arguments would require a more sophisticated model; one that would allow for a trade-off between the cost of monitoring and the quality of the observed signal.

The following result adds yet another reason in favor of contest schemes; one that justifies their popularity and explains their advantage even in the cases of independence and complementarities across agents. The result extends corollary (5) to the case of multi-agent incentive provision. It shows that if the principal has a limited budget to spend on incentives, a tournament is the most effective way of motivating the team. A little more structure is required before proceeding; in particular, the model so far allows for one agent's effort to influence more another agent's signal than her own. The following condition is therefore pretty intuitive.

Condition 15 (low externalities) *For all* $x \in [0, \bar{x}]^n$, *all* $a \in A^n$, *and all* i; $\frac{\partial}{\partial x_i}\lambda_i(x,a) > \sum_{j \neq i} \left| \frac{\partial}{\partial x_j}\lambda_i(x,a) \right|$

The condition implies that externalities are never too big; observing a better signal from i is a better indicator of i's effort choice than observing all other signals moving in the direction of i's externalities.

The problem that the principal faces can be expressed as follows:

$$\begin{gathered}\min_{\{s_i\}_i} \sum_i E[s_i(x) | a^*] \\ \text{s.t.} \quad \forall i \quad (4) \\ \forall i \quad s_i(x) \geq L \quad (LL_i) \\ \sum_i s_i(x) \leq B \quad (LL_P).\end{gathered} \tag{9}$$

Where LL_i denotes agent i's limited liability constraint, and LL_P the principal's liability constraint.

Proposition 16 *Suppose the principal cannot make payments in excess of B in any state of the world. Then, the optimal incentive scheme for a team of symmetric risk-neutral agents with low positive externalities takes the form of a tournament with a standard, and a single prize (winner takes all).*

Proof. See appendix ■

Proposition 11 showed that the least costly way of motivating a team of agents with positive externalities is through a joint super-bonus. Here, I show that the correct approximation to such a scheme relies on giving the full bonus to a single agent: a tournament. Levin (2002) finds a similar result when considering dynamic relational contracts. In his model, the principal has not a limited budget, but a similar constraint arises from the fact that parties may run away, and no one, the principal in particular, is willing to pay more than the value of future interaction.

4 Interim monitoring

The previous section investigated the form of minimum cost incentive schemes in a static relationship. This section addresses the question of how robust such incentive schemes are to environments where agents can adapt to partial revelation of information.

Suppose agents' actions are not taken once and for all, but need to be taken on a daily basis. If agents have a rough idea about their interim production, they will adapt their actions according to the updated information. The IPE, JPE, and tournament incentive schemes that resulted from the analysis in the previous section need no longer be optimal. Solving for the optimal incentive scheme under these circumstances is left for future research. In this section I will address the positive question of how agents react to such incentive schemes when allowed to update their actions.

Consider a car seller who gets a bonus for attaining a monthly sales quota. Given that the seller learns throughout the month how close he is to the quota, he will work harder when he is close to the quota, and less hard when either the quota is nearly impossible to reach, or when he has already reached it. Similarly, in a golf tournament, the leader may put more effort when competition is tight than when he has a comfortable gap over the second player (Ehrenberg and Bognanno (1990) shows some evidence of such behavior).

Most of the results in this section will be based on the following simplified extension of the model. There are only two ex-ante symmetric agents. Before investing in effort, agents observe an interim measure of performance $y =$

(y_1, y_2). The measure can be interpreted as output from an initial period, but I will not model first period's effort choices. Instead, the signal y is assumed to be distributed according to an exogenous (symmetric) probability distribution $G(\cdot)$, with density $g(\cdot)$, and support in $[0, \bar{y}]^2$. Final output will be assumed to behave according to $x_i = y_i + z_i$, where z_i is distributed according to $F^i(\cdot|a)$,with support on $[0, \bar{z}]$, density $f^i(\cdot|a)$, and $a = (a_1, a_2)$ is the vector of efforts. The productivity and independence conditions will be assumed to hold throughout. The function $a_i(y)$ will denote the equilibrium action of agent i when observing y, and $a(y) = (a_1(y), a_2(y))$.

The rest of the section is divided into five parts. In parts 1-3, I start by considering agents' response under three type of incentive schemes: i) individual performance evaluations (IPE), where agents earn individual bonuses for reaching a threshold level of production, ii) joint performance evaluation (JPE), where a common bonus is given when aggregate production $(x_1 + x_2)$ reaches some target, and iii) tournaments, where only the agent with the highest output may be rewarded. Part 4 summarizes the results of the previous parts to draw some conclusions regarding compensation schemes and correlation of production. Finally, part 5 extends the analysis of tournaments to examine strategic considerations.

4.1 Individual Performance Evaluation

The car seller and the golfer mentioned above work harder when their chances of obtaining the prize are fair, rather than very high or low. It is when their effort has the highest impact on their expected payoffs that they will go the extra mile. This subsection shows that such behavior is optimal when agents face an individual bonus scheme. The following two subsections take care of the joint bonus and tournament cases.

Suppose agent i 's payoff has the following form

$$s_i(x) = \begin{cases} 1 & x_i \geq M \\ 0 & x_i < M \end{cases}.$$

I will say that agent i is *on schedule* (given y) if $E[z_i|a(y)] = M - y_i$. That is, if the expected output in equilibrium is exactly the bonus threshold level. When the equality is changed for $>$ $(<)$, it will be said that the agent is above (below) schedule. If the above intuition is correct, an agent would exert the highest effort when on schedule. The following reasonable condition obtains this result.

Condition 17 (Effort Intensity)

For every a, $\left(-\frac{\partial}{\partial a_i}F^i(\cdot|a)\right)$ *is increasing on* $[0, z_i(a)]$ *and decreasing on* $[z_i(a), \bar{z}]$*, where* $z_i(a) := E[z_i|a]$.

The condition says that effort is most effective at the mean of production. That is, the marginal (probabilistic) increase in production due to effort is more sensitive around the mean that anywhere else. The condition is satisfied, for example if $z_i = a_i + \varepsilon_i$, where ε_i is a noisy term with zero mean and a distribution with a single peak at zero. The following proposition characterizes the equilibrium response of the agents to interim information under a bonus scheme. It shows that agents' dedication depends on how close to the schedule they are performing.

Proposition 18 *Suppose the Productivity and Effort Intensity conditions hold. If for all* a, $F^i_{a_i a_i}(\cdot|a) \leq c''(a_i)$*, then in any Nash equilibrium* $(a_1(y), a_2(y))$*,* a_i *is non-decreasing (non-increasing) in* y_i*, whenever agent* i *is below (above) schedule.*

Proof. Consider a Nash equilibrium $a(y)$, and a point $y = (y_1, y_2)$ where agent 1 is below schedule. That is, where $z_1(a(y)) = E[z_1|a(y)] < M - y_1$. From the intensity of effort condition it follows that $\left(-\frac{\partial}{\partial a_1}F^1(\cdot|a)\right)$ is decreasing around $M - y_1$. Fix a level of effort a_2 for agent 2, and consider 1's objective function:

$$\left(1 - F^1(M - y_1|a, a_2)\right) - c(a).$$

Let $a(y_1, a_2)$ denote the maximizer.[4] Letting $a_2 = a_2(y)$, it follows that the objective function has increasing differences[5] in (a, y_1) around y. Hence, by Topkis (1998, theorem 2.8.1 p. 76), it must be that the maximizer a is non decreasing in y_1.
The proof for the case where the agent is above schedule is identical. ■

The key assumption for the proposition to hold is the Effort Intensity condition. It implies that the marginal increase in the probability of attaining a given level of output is the highest when such output is precisely at the mean. In this way, it gives more structure to the model in terms of *how* is effort affecting output.

[4] Given that $F^i_{a_i a_i}(\cdot|a) \leq c''(a_i)$, the maximizer is unique

[5] With differentiable functions of real variables, $f(x, y)$ has increasing differences in (x, y) if $f_{xy} \geq 0$. A general discussion can be found in Topkis 1998.

4.2 Contests

Now consider the case of a tournament. To make things simple, suppose there is no standard; hence agents compete only against each other. After observing a signal $y = (y_1, y_2)$, assuming agent 2 exerts effort a_2, agent 1 will choose a as to maximize

$$\Pr\{z_1 - z_2 > y_2 - y_1 | a, a_2\} - c(a).$$

As in the IPE case, the prize is normalized to 1. Let $H(\cdot|a)$ denote the c.d.f. of $(z_1 - z_2)$ given effort $a = (a_1, a_2)$. A similar condition to effort intensity makes it possible to characterize the nature of competition given the observed state. I shall use a slightly more restrictive condition that will yield a stronger result.

Condition 19 (Effort Gap) *Let $d = a_1 - a_2$, and $\Delta(d) = E[z_1 - z_2|d]$, then*
i) $H(\cdot|a_1, a_2) = \bar{H}(\cdot|a_1 - a_2)$,
ii) $d \geq d'$ implies $\bar{H}(\cdot|d) \succeq_{LR} \bar{H}(\cdot|d')$,
iii) For every d; $-\bar{H}_d(\Delta|d)$ is increasing for $\Delta \leq \Delta(d)$, and decreasing thereafter, and
iv) $|H_{dd}| \leq \inf_{a_i \in A} c''(a_i)$.

The condition is satisfied, for example, if each agent controls the mean of a Normal distribution with the same fixed variance. The first and third parts are the essence of the condition; they restricts the way that efforts may affect outputs. Part i) implies that the probability distribution of the gap in second period's outputs, $z_1 - z_2$, depends only on the difference of efforts between agents 1 and 2. Part iii) accomplishes in tournaments the same goal that condition Effort Intensity did in IPE's; it implies that effort affects more the probability distribution around the expected production gap. Part ii) says that the effort difference orders the distribution over the gap in the likelihood ratio order sense. Finally, part iv) guarantees that each agent's problem is concave.

As in the IPE case, the effort gap condition yields a result that confirms the intuition about parties competing more fiercely when none has a significant advantage. In this setting it turns out that, in equilibrium, both parties will exert the same amount of effort always. That is, in equilibrium, the leader's effort to maintain her advantage is the same as the laggard's effort to catch-up.

Proposition 20 *Suppose the Effort Gap condition holds. Then, for any $y = (y_1, y_2)$ there is a unique Nash Equilibrium in Pure Strategies $(a_1(y), a_2(y))$,* *and $a_1(y) = a_2(y)$ for all y. Moreover, equilibrium effort is a decreasing function of $|y_1 - y_2|$.*

Proof. Given $\Delta = y_2 - y_1$, the objective functions for agents 1 and 2 are, respectively:

$$\begin{aligned}&\left[1 - \bar{H}(\Delta|a_1 - a_2)\right] - c(a_1),\\ &\bar{H}(\Delta|a_1 - a_2) - c(a_2).\end{aligned}$$

Each agent takes the other agent's action as given. In equilibrium then, (a_1, a_2) must simultaneously solve the first order conditions:

$$\begin{aligned}-\bar{H}_d(\Delta|a_1 - a_2) &= c'(a_1)\\ -\bar{H}_d(\Delta|a_1 - a_2) &= c'(a_2)\end{aligned}$$

It follows that $c'(a_1) = c'(a_2)$, and since $c'' > 0$, $a_1 = a_2$. Note that the level of effort is always interior. Now, the value of $a(y) = a_1(y) = a_2(y)$ is determined by

$$c'(a(y)) = -H_d(y_2 - y_1|0).$$

Therefore, given y, there is only one possible equilibrium level of effort.
By symmetry $\Delta(0) = 0$. Part iii) of the effort gap condition implies then that effort is highest when $y_1 = y_2$, and decreases in $|y_1 - y_2|$. ■

The proposition offers an easy benchmark to analyze different extensions. For example, one may ask how tournament participants react when they face learning-by-doing (LBD), or decreasing returns to effort (DRE) effects. Suppose the cost of effort for agent i depends on her accumulated production y_i. If $c_y \leq 0$ (LBD), it should be the case that the leader will exert more effort than the laggard. On the contrary, if $c_y \geq 0$ (DRE) it is the laggard who works harder to try to catch-up.

4.3 Joint performance compensation

Consider now the case where both agents receive a bonus when their joint production reaches a given threshold M. When observing signals (y_1, y_2), agent 1 will try to maximize

$$\Pr\{z_1 + z_2 > M - (y_1 + y_2)|a, a_2\} - c(a),$$

where a_2 is the equilibrium action of agent 2. The result, as you could have anticipated, is similar to the previous cases: Agents exert more effort when close to the schedule; and both agents will work equally hard –unless one would incorporate LBD or DRE effects.

Let now $H(\cdot|a)$ denote the c.d.f. of $s = z_1 + z_2$; the relevant condition for this case can be stated as follows.

Condition 21 (Effort Sum) $H(\cdot|a_1, a_2) = \bar{H}(\cdot|a_1 + a_2)$. *Denoting* $e = a_1 + a_2$; $s(e) = E[s|e]$, *then:*
i) $e > \hat{e}$ *implies* $\bar{H}(\cdot|e) \succeq_{LR} \bar{H}(\cdot|\hat{e})$,
ii) For every s, $-\bar{H}_e(s|e)$ *is increasing for* $s \leq s(e)$, *and decreasing thereafter*
iii) $\left|\bar{H}_{ee}(s|e)\right| \leq \inf_{a \in A} c''(a)$ *for all* (s, e)

The interpretation of the condition is the same as the conditions used for the previous cases. The key feature is that the probabilistic contribution of effort to output is highest at the expected output level. As in the IPE case, agents are said to be below (above) schedule if

$$E[s|a(y)] < (>)\, M - y_1 - y_2.$$

Proposition 22 *Suppose condition Effort Sum holds. Then, under a JPE scheme, there are symmetric equilibria for which i) both agents exert the same effort and ii) effort is increasing (decreasing) in aggregate production* $(y_1 + y_2)$ *when output is below (above) schedule.*

Proof. Let $Y = y_1 + y_2$. Assuming player j chooses effort a_j, player i solves

$$\max_a 1 - \bar{H}(M - Y|a + a_j) - c(a). \tag{10}$$

Thus, a Nash equilibrium (a_1, a_2) must solve simultaneously the first order conditions

$$\begin{aligned} \bar{H}_e(M - Y|a_1 + a_2) &= c(a_1) \text{ and} \\ \bar{H}_e(M - Y|a_1 + a_2) &= c(a_2). \end{aligned}$$

This implies $a_1 = a_2$. Now, fixing the level of a_j, the objective function (10) has increasing differences in (a, Y) when below schedule. The result follows from Topkis (1998, theorem 2.8.1). ■

The proof of the proposition is similar to that of propositions (18) and (20) is omitted.

4.4 Choice of Compensation Scheme and Correlation

This subsection argues that, to a certain extent, the choice of compensation scheme is driven by the correlation of agents' initial output (y_1, y_2). This observation is derived directly from the previous results of the section (propositions 18, 20, and 22), and the discussion is informal.

Suppose the principal is interested in the sum of second period efforts. It is natural to ask when $a_1(y) + a_2(y) \geq a^*$, for a given a^*. The results so far suggest a different answer for each incentive scheme. First, under an IPE scheme, the region (in the y-space) has the shape of a diamond around the point (y^s, y^s), where y^s is the on schedule production level. Second, under a JPE scheme, the region is a neighborhood of the line $y_1 + y_2 = 2y^s$. Third, under a tournament scheme, the region is a neighborhood of the line $y_1 = y_2$. See Figure 1.

The interesting feature of such regions is that they suggest a rule of thumb for designing incentives. If agents' interim outputs (y) are likely to be positively correlated, then a tournament will generate higher levels of effort more often. Whereas if interim outputs are likely to be negatively correlated, rewarding joint performance is more productive. In other words, tournaments will perform better when the tasks of members of the team are complements, and joint bonus scheme should be preferred when members are partial substitutes.

Lazear and Rosen (1981), Holmstrom (1982), Green and Stokey (1983), and Nalebuff and Stiglitz (1983a) all make the point that there is a benefit in using tournaments when agents' output is positively correlated. The same conclusion arises here, but for a very different reason. Whereas they point out that tournaments filter shocks that affect all agents similarly, the role of correlation here is to keep competition among agents close. And close competition generates higher effort levels when agents can verify the state of the tournament—Proposition 20.

A question that arises from this observation and is left for future research is the following. If the compensation scheme affects agents interactions, one would expect to see cooperative behavior under JPE's, and little cooperation under tournaments. Suppose the space of actions is richer so that agents can choose how much to cooperate. Even though a tournament seems to be a better scheme under complementarities, the fact of putting in place the tournament should reduce, if not eliminate, such complementarities. Similarly, under a JPE scheme, substitute agents may start to cooperate and help each other up to the point that they become complements.

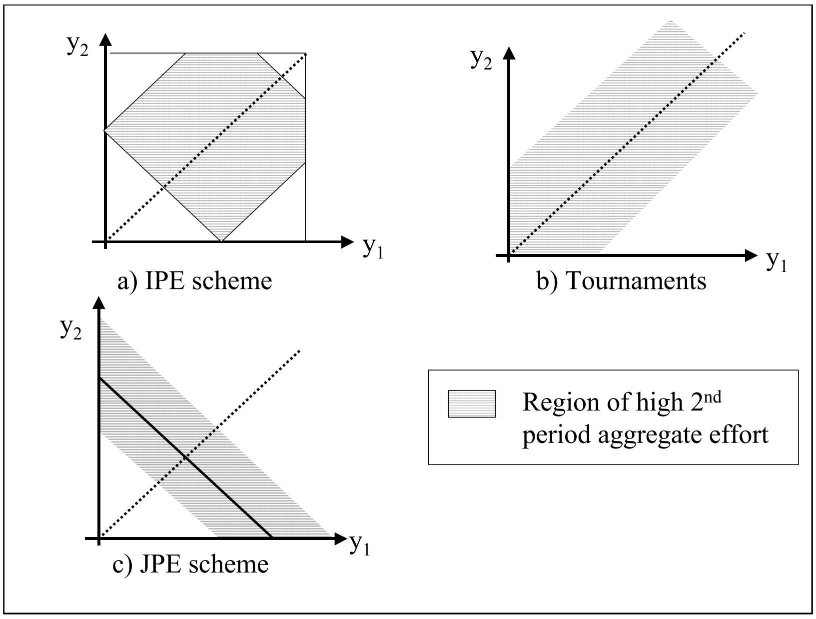

Figure 1: Correlation and Incentives

4.5 Strategic Considerations in Tournaments

The model set so far also allows to draw some conclusions regarding the players strategic dynamic interaction. Following the terminology of Bulow, Genakoplos and Klemperer (1985), when should one expect players to see their efforts as *strategic substitutes* or *complements*? That is, if agents could monitor each other perfectly, would more aggressive play from one agent—i.e., choosing effort above the equilibrium level—induce the other to increase or decrease her choice of effort? Clearly, propositions 18, 20 and 22 imply that optimal actions are not monotonic with respect to revealed information, but how are best responses affected by a deviation from the oponent?

I analyze this question by restricting attention to the case where the agent controls the mean of production and noise is independent of the action. That is, $z_i = a_i + \varepsilon_i$, where the ε_i's are i.i.d.[6] zero mean random variables, with support on $(-\infty, \infty)$.[7]

Agent i is said to regard agent j as a *strategic partner (competitor)* if i's best response to an increase in j's effort choice is increasing (decreasing)

[6] i.i.d. stands for independent and identically distributed.

[7] Assuming full support is not necessary, but avoids consideration of the cases where the first period determines completely the ranking.

around equilibrium efforts.[8] Consider the problem of agent 1. After observing $\Delta = y_2 - y_1$, she will choose a_1 as to maximize

$$\Pr\{\varepsilon_1 - \varepsilon_2 \geq \Delta + a_2 - a_1\} - c(a_1).$$

If $\hat{H}(\cdot)$ denotes the c.d.f. of $\varepsilon_1 - \varepsilon_2$, condition effort gap and i.i.d. errors, imply that $\hat{H}$ is symmetrically distributed around zero, and its density, $\hat{h}$, has a single peak at zero. The first order condition for agent 1 can be written as

$$\hat{h}(\Delta + a_2 - a_1) = c'(a_1). \tag{11}$$

From proposition 20 it follows that equilibrium efforts are $a_1^* = a_2^* = \psi\left(\hat{h}(\Delta)\right)$, where $\psi(\cdot)$ is the inverse of the marginal cost function, that is

$$\psi(c) = a \Leftrightarrow c'(a) = c.$$

Inspection of the first order condition shows that if $\Delta < 0$—agent 1 is the leader—an upwards deviation in 2's effort will drive the left hand side of (11) up, and agent 1's best response would be to increase her level of effort. Similarly, if $\Delta > 0$—agent 1 is the laggard—, an upward deviation from 2 will cause 1's best response to decrease. The result is illustrated in figure 2, and summarized in the following proposition.

Proposition 23 *Suppose the Effort Gap condition holds, and that $z_i = a_i + \varepsilon_i$, with ε_1 and ε_2 independent zero mean random variables with support on $(-\infty, \infty)$. Then the leader in a tournament will regard her adversary as a strategic partner, whereas the laggard will regard hers as a strategic competitor. That is, the leader's (laggard's) best-response correspondence is upward (downward) sloping around equilibrium.*

The importance of the proposition comes from its implications to settings where agents observe each other more closely. The result suggests that i) by deviating upwards, the leader induces the laggard to reduce her effort, therefore increasing the expected gap between them. ii) If it is the laggard who attempts to catch-up faster with the leader by deviating upwards, the leader will respond by increasing her effort, protecting her lead.

Dixit (1987) shows that, if possible, the advantaged player would commit to exert more effort than she would exert in equilibrium. Consider a Stackelberg version of the game. Suppose player 1 is the leader and chooses effort

[8] Note that, as stated, the condition only looks at strategic complementarities (substitutes) locally.

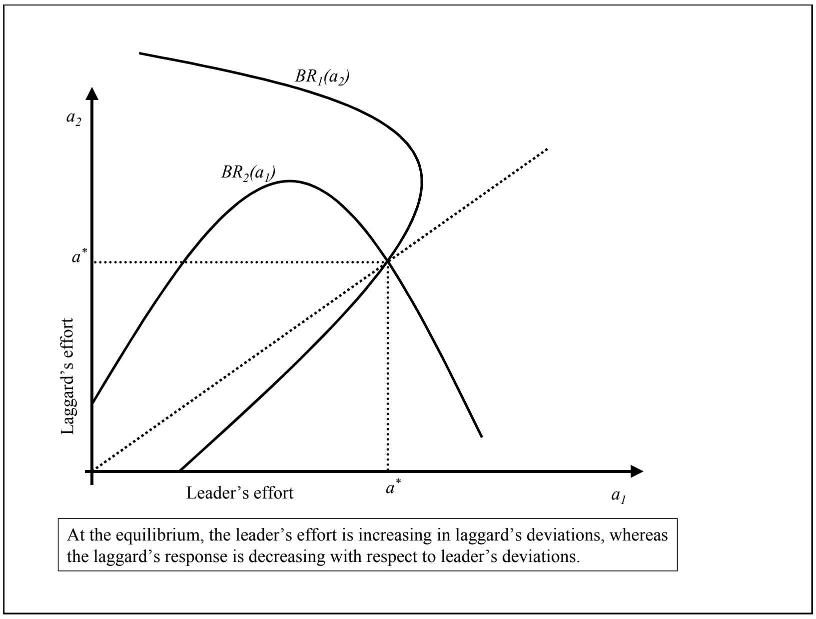

Figure 2: Best Response Correspondences

before player 2. Suppose also that player 2 observes her opponent's choice. In this game, player 1 selects a point on 2's best response correspondence. Incorporating the fact that player 1 affects 2's choice, the first order condition analogous to (11) can be written as

$$\hat{h}\left(\Delta + a_2\left(a_1\right) - a_1\right) - a_2'\left(a_1\right)\hat{h}\left(\Delta + a_2\left(a_1\right) - a_1\right) = c'\left(a_1\right).$$

The second term on the left represents the strategic effect of player 1 's choice of effort. If player 1 has an advantage ($\Delta < 0$) then $a_2' \leq 0$. Hence $c'\left(a_1\right) \geq \hat{h}\left(\Delta + a_2\left(a_1\right) - a_1\right)$. A similar point can be made when player 1 is the laggard. In summary, if the leader (laggard) moves first, she will exert more (less) effort than if both players move simultaneously.

Starting from a symmetric position, agents may then have an incentive to jump-start, gaining a small advantage initially in the competition. This point is analyzed below by comparing equilibrium levels of effort with and without revelation of information.

4.6 Anticipation and the Role of Information in Tournaments

Consider an extended model that includes effort in the first period—i.e., the distribution of y also depends on a costly action. Several comparative static results emerge. For example: when does one expect agents to start aggressively and slow down later? When will they work harder under no information?

Suppose now that y responds to agents' effort in a similar manner as z does. Specifically, let $e = (e_1, e_2)$ denote first period's effort, and

$$y_i = e_i + \eta_i. \tag{12}$$

Where the η_i's are i.i.d. random variables that have the same distribution as the ε_i's. Suppose also that effort is equally costly in each period. In other words, there are two identical periods, and output is additive across periods.

First consider the case of no revealed information after the first period—that is, agents take decisions about a without knowing y. The effort decision of each agent can be viewed as a single decision. Letting $\nu_i = \eta_i + \varepsilon_i$, agent's i objective can be written as:

$$\Pr\{\nu_i - \nu_j \geq e_j - e_i + a_j - a_i\} - c(e_i) - c(a_i).$$

Since $c'' > 0$, it follows that each agent i will find in their benefit to set $e_i = a_i$. Note that $(\nu_1 - \nu_2)$ has the same distribution as $2(\varepsilon_1 - \varepsilon_2)$. Fixing $s_2 = a_2 + e_2$, the problem for agent 1 is to choose s, as to maximize

$$1 - \hat{H}\left(\frac{s_2 - s}{2}\right) - 2c(s/2).$$

Where, again, $\hat{H}$ is the c.d.f. of $(\varepsilon_1 - \varepsilon_2)$.It is readily seen that there is a symmetric Nash equilibrium where agents chose $e_i = a_i = \psi\left(\frac{\hat{h}(0)}{2}\right)$, for $i = 1, 2$. Assume the equilibrium is unique.[9]

Consider now the case with interim revelation of information. It follows from (11) that in the second-period's symmetric Nash equilibrium, effort is given by $a_1(\Delta) = a_2(\Delta) = \psi\left(\hat{h}(\Delta)\right)$, where $\Delta = y_2 - y_1$.

In the first period, agent 1's objective is to maximize the expected value of the probability of obtaining the prize minus her cost of effort. Given that, regardless what happens in the first period, $a_1 = a_2$ in the second period;

[9] A sufficient condition would be that $\left|\hat{h}'\right| < \inf_a 2c''(a)$.

the probability of winning the prize for agent 1 is $1 - \hat{H}(\Delta)$. The objective for agents 1 and 2 can thus be expressed, respectively, as

$$E\left[1 - \hat{H}(\Delta) - c(a_1(\Delta))\right] - c(e_1),$$
$$E\left[\hat{H}(\Delta) - c(a_2(\Delta))\right] - c(e_2).$$

Where expectations are taken with respect to Δ. Noting that $\Delta = \eta_2 - \eta_1 + e_2 - e_1$; the first order conditions that each agent solves are

$$\begin{aligned} c'(e_1) &= E\left[\hat{h}(\Delta)\right] + E\left[a_1'(\Delta)\hat{h}(\Delta)\right], \\ c'(e_2) &= E\left[\hat{h}(\Delta)\right] - E\left[a_2'(\Delta)\hat{h}(\Delta)\right]. \end{aligned}$$

It follows from the symmetry assumptions of the problem that the last term on the right of each equation is zero. To see this, note that $\hat{h}$ is symmetric around zero and so is $a_i(\cdot)$. Now, since $a_i(\cdot)$ is symmetric, then it must be that $a_i'(\Delta) = -a_i'(-\Delta)$, and the whole expected value—which is taken w.r.t. a symmetric distribution—cancels out.

Summarizing, first period efforts when information is available are given by

$$e_i = \psi\left(E\left[\hat{h}(\Delta)\right]\right). \tag{13}$$

Expected second period efforts satisfy:

$$E[a_i] = E\left[\psi\left(\hat{h}(\Delta)\right)\right]. \tag{14}$$

Whereas with no information, first and second period efforts are equal and are given by:

$$\bar{e}_i = \bar{a}_i = \psi\left(\frac{\hat{h}(0)}{2}\right). \tag{15}$$

Jensen's inequality provides a direct comparison between first and second period efforts when information is available. The condition relies on the convexity of marginal costs, that is on the sign of c'''. For example, if c' is convex, then first period efforts are higher than expected effort in the second period. Similar to the notion of prudence in Kimball (1990), the sign of c''' here determines how sensitive is the effort decision to risk.

Comparisons between equilibrium efforts with and without information depend on the nature of uncertainty. Higher efforts in the no information case require that the relative likelihood of observing very small gaps be high. That is, if the effect of noise is small, there is good chance that competition will be tight after the first period, and hiding information from agents may procure better incentives. The results are summarized below.

Proposition 24 *Assume the conditions of Proposition 23, and that first period output follows* (12) *. Then, for* $i = 1, 2$
a) If c' *is convex (concave) then* $e_i \geq (\leq) E[a_i]$ *.*
b) If $E\left[\hat{h}(\Delta)\right] \geq \frac{\hat{h}(0)}{2}$ *(i.e. noise is spread out) then* $e_i \geq \bar{e}_i$*.*
c) If $E\left[\hat{h}(\Delta)\right] \leq \frac{\hat{h}(0)}{2}$ *(i.e. uncertainty is concentrated) then* $e_i \leq \bar{e}_i$*.*
d) If c' *is convex and uncertainty concentrated, total effort under no information is higher than expected total effort under information.*
e) If c' *is concave and uncertainty spread out, total effort under no information is lower than expected total effort under information.*
Where e_i *and* $\bar{e}_i$ *denote first period effort of agent* i *in the information, and no information cases respectively. And* a_i *is second period effort of agent* i *in the information case.*

Proof. a) follows from Jensen's inequality applied to (13) or (14). A direct comparison of 13 and 15 imply both b) and c). Finally, d) and e) are consequences of a), b), and c). ■

It is hard to convey an economic intuition for the relevance of the sign of c''' in these comparisons. One explanation is that if the marginal cost of effort increases very rapidly (c' convex), then it pays to anticipate and exert high efforts in the first period. It would be too costly to try to recover in the second period. Similarly, if the marginal cost of effort does not increase so rapidly, agents can afford to wait until the second period, and need not be so preemptive initially.

5 Concluding Remarks

This paper addresses three points that enhance our understanding of a firm's choice of compensation scheme. First, it derives the form of contracts that minimize the cost of providing incentives in static relationships. Second, common institutional constraints are associated with different forms of optimal incentive contracts. Third, the paper investigates how agents adapt their response to optimal (static) contracts when they have access to interim information.

I show that different work environments call for different forms of compensation schemes. With no restrictions on the contract space, large bonuses that reward only the best possible performance minimize the cost of providing incentives. Although these contracts would not be possible in practice (they require infinite payments), approximate versions are found in industries like sports, music and movies: Top performers receive multi-million dollar sums, whereas average performances may not be rewarded at all. Less extreme approximations are found in less popular industries. For example, Zenger (1992) studying compensation plans in Silicon Valley companies finds that economic incentives for engineers depend on extreme performances: awarding bonuses for top performers, and laying-off under-performers. Corollaries 5 and 6 show that for a principal that faces a limited liability condition, optimal contracts with a single agent are similar to those observed by Zenger. Different conditions result in other commonly used incentive schemes. A monotonicity condition (on the shares that go to both principal and agent) gives rise to stock-option plans (Corollary 7). Group rewards are shown to be effective under team production complementarities (Proposition 11). However, the most surprising result is the optimality of tournaments under weak conditions: Proposition 16 shows that with double sided limited liability, regardless of production externalities across agents—positive, negative or none—single prize tournaments are the most efficient way to motivate a team.

There is a myriad of reasons for why a particular incentive scheme that is optimal under a simple representation of reality does not perform well in practice. The multi-task problem introduced by Holmstrom and Milgrom (1991) alerts us to the fact that as long as contracts fail to include all possible contingencies, dysfunctional behavior is likely to arise in equilibrium. Contracts are incomplete. Many contingencies are impossible to foresee or to contract for; different parties to a contract will seek to use them to their advantage as events unfold. Therefore, the simple conclusions about choice of compensation schemes derived from the static model can not tell a complete story. They are useful, no doubt, as long as they be seen as servants and not masters of contract design.

The paper addresses one dimension of contract incomplteness by studying agents' response to interim information. Rather than solving for optimal contracts in this setting (a topic for future research), this study characterizes agents' behavior under bonus and tournament schemes, and shows conditions that favor one incentive scheme over another. Propositions 18, 20 and 22 describe the type of dysfunctional behavior that optimal (static) schemes may induce when agents learn about their performance. Agents exert low levels of effort when the chance of obtaining the bonus (or prize in a tournament) becomes either very small, or very high. However, positive results also emerge from this last part of the analysis. A positive correlation in agents' production favors tournament-type schemes over bonuses. The reason for this is not the standardly assumed one, that tournaments filter out common shocks. Rather, because outputs are possitively correlated, there is a better chance that competition will be close when information is available, and close competition induces more effort in tournaments. Similarly, a negative correlation in outputs favors group rewards because aggregate production is likely to be average, and a joint bonus scheme would induce low levels of effort when observing very high or very low production signals.

The final results illustrate the type of interaction that one may expect under a promotion tournament. With only two agents and no production externalities, Proposition 23 characterizes the relationship between the leader and the laggard. The former sees her competitor as a strategic partner; meaning that the harder the laggard works, the harder she will work also. On the other hand the laggard has the opposite viewpoint. She regards the leader as a strategic substitute; that is, if she observes the leader working harder than equilibrium would predict, she would respond by slowing down. Similar effects are present in R&D races; a careful analysis of the dynamics of interaction in tournaments is likely to yield promising results, and is left for future research. Proposition 24 shows that tournaments unravel in different ways, the reasons for which are quite subtle. First, whether participants prefer to start slowly, waiting for information before they adjust their pace, or whether they would rather act preemptively, trying to win an early lead, depends on the convexity of marginal costs of effort. Second, the same proposition sheds some light about the role of information in providing incentives. By comparing agents' behavior when they have access to interim information with the case when they do not, it is shown that the degree of uncertainty determines the level of incentives. Access to information is better (induces more effort) when output signals are noisier; that is, if chance plays an important role in determining outcomes, agents have stronger incentives when they will be able to know the state of the tournament halfway through.

6 Appendix

Proof. (of proposition 1)It is easy to see that any contract that satisfies inequalities (ii)' and (iii), and satisfies (i) with equality is a solution to P. It suffices then to choose appropriate values for $\tilde{x}$, M and k. I will choose such values as to make sure that (i) binds and (ii)' is satisfied, and then show that if condition 2 holds then (iii) must also hold.

Since $\frac{f_a}{f}$ is continuous and increasing, there is some $\tilde{x}$ such that $f_a\left(\tilde{x}|a\right) >$

0 and such that, for all $x > \tilde{x}$,

$$\bar{u} + c(a) > L + c'(a) \frac{f(x|a)}{f_a(x|a)}. \tag{16}$$

Now define k as

$$k := \frac{c'(a)}{\int_{\tilde{x}}^{\bar{x}} f_a(x|a)\, dx}.$$

Finally, define M as

$$M = \bar{u} + c(a) - k(1 - F(\tilde{x}|a)).$$

By construction, $s(x) = M + k1_{x \geq \tilde{x}}$ satisfies (i) with equality, and (ii)'; it remains to show that it also satisfies (iii) –i.e. that $M \geq L$.From 16 and the definition of M, (iii) will be satisfied whenever $k(1 - F(\tilde{x}|a)) < c'(a) \frac{f(\tilde{x}|a)}{f_a(\tilde{x}|a)}$. The proof is completed by noting that:

$$\begin{aligned} k(1 - F(\tilde{x}|a)) &= \frac{c'(a)(1 - F(\tilde{x}|a))}{\int_{\tilde{x}}^{\bar{x}} f_a(x|a)\, dx} \\ &= \frac{c'(a)(1 - F(\tilde{x}|a))}{\int_{\tilde{x}}^{\bar{x}} \frac{f_a(x|a)}{f(x|a)} f(x|a)\, dx} \\ &= \frac{c'(a)}{E\left[\frac{f_a(x|a)}{f(x|a)}|a, x \geq \tilde{x}\right]} \\ &\geq c'(a) \frac{f(\bar{x}|a)}{f_a(\bar{x}|a)}, \end{aligned}$$

where the inequality follows from MLRP.

Note that as $\tilde{x} \uparrow \bar{x}$, $(1 - F(\tilde{x}|a)) \downarrow 0$. However, by construction k depends on $\tilde{x}$ and $k(1 - F(\tilde{x}|a)) \downarrow c'(a) \frac{f_a(\bar{x}|a)}{f(\bar{x}|a)}$ (monotonicity is due to the MLRP assumption.) This suggests that the condition is also necessary, as proposition 3 shows. ■

The proof of proposition 3 is presented in two steps. First I show that any solution must pay the minimum possible transfer for low levels of output. Next, using the first step, the program is transformed as to derive a characterization of the solution in the high levels of output region.

Claim 25 *For almost every x in $A = \{x : f_a(x|a) < 0\}$, any solution to problem RP satisfies $s(x) = L$.*

Proof. By way of contradiction. Suppose that $s(\cdot)$ is a solution with $P\{x \in A : s(x) > L\} > 0$.
From the likelihood ratio ordering, there must be some $\hat{x}$ such that $f_a(x|a) > 0$ iff $x > \hat{x}$. Then $d := -\int_0^{\hat{x}} [s(x) - L] f_a(x|a)\, dx > 0$. Noting that $\int_0^{\bar{x}} f_a(x|a)\, dx = 0$, it follows from (ii)' that

$$\int_{\hat{x}}^{\bar{x}} [s(x) - L] f_a(x|a)\, dx = c'(a) + d.$$

Define

$$\hat{s}(x) = \begin{cases} L + (s(x) - L) \frac{c'(a)}{c'(a)+d} & x \notin A \\ L & x \in A \end{cases}$$

It follows that $s \geq \hat{s}$ with $E[s(x)|a] > E[\hat{s}(x)|a]$, and $\hat{s}$ satisfies constraints (ii)' and (iii). This contradicts the assumption that s is a solution. ■

It remains to characterize the solution for $x > \hat{x}$.

Proof. (of proposition 3) Given the previous claim, the program can be rewritten as:

$$\begin{aligned} &\min_{s(x)} \int_{\hat{x}}^{\bar{x}} (s(x) - L) f(x|a)\, dx \\ \text{s.t.} \quad &\int_{\hat{x}}^{\bar{x}} (s(x) - L) f_a(x|a)\, dx = c'(a) \\ &s(x) - L \geq 0 \end{aligned}$$

And since $f_a(x|a) > 0$ for $x > \hat{x}$, the change of variable $g(x) := (s(x) - L) f_a(x|a)$ yields the equivalent program

$$\begin{aligned} &\min_{s(x)} \int_{\hat{x}}^{\bar{x}} g(x) \frac{f(x|a)}{f_a(x|a)} dx \\ \text{s.t.} \quad &\int_{\hat{x}}^{\bar{x}} g(x)\, dx = c'(a) \\ &g(x) \geq 0 \end{aligned} \tag{17}$$

From the likelihood ratio ordering it follows that $\frac{f(x|a)}{f_a(x|a)}$ is decreasing in x. The program then has a "trivial" solution: one wants to put all the weight of g at the extreme point where $\frac{f(x|a)}{f_a(x|a)}$ is minimum. Hence, with no further constraints the solution is to use an impulse function with a mass point of size $c'(a)$ at $\bar{x}$. That is, $g(x) = c'(a)\,\delta(x - \bar{x})$ is the solution. That is, $s(x) = L + c'(a)\,\delta(x - \bar{x}) / f_a(x|a)$. ■

Proof. (of proposition 11) The principal is minimizing the expected value of the sum of payments $E\left[\sum S_i(x)|a^*\right]$. She can achieve this objective by minimizing each agent's expected payment individually. Define $g(x) = (S_i(x) - L_i)\,\mu(x|a^*)$, problem (5) is reduced to solving:

$$\min_{g(\cdot)} \int g(x)\,dx$$

$$\text{subject to}$$

$$\int g(x)\left[\sum \frac{f^j_{a_i}\left(x_j|a_i^*, a_{-i}^*\right)}{f^j\left(x_j|a_i^*, a_{-i}^*\right)}\right] dx_1...dx_n = c_i'(a_i^*)$$

$$g(x) \geq 0$$

As in proposition 3, given no further constraints on the function $g(\cdot)$, the optimum is obtained by concentrating all the weight of $g(\cdot)$ where $\lambda_i(x, a^*) = \sum \frac{f^j_{a_i}(x_j|a^*)}{f^j(x_j|a^*)}$ is maximum, that is when $x \in \arg\max_x \lambda_i(x, a^*)$.
With positive externalities, $\bar{\mathbf{x}} = (\bar{x}_1, ..., \bar{x}_n) \in \arg\max_x \lambda_i(x, a^*)$, whereas with negative externalities $(\bar{x}_i, 0) \in \arg\max_x \lambda_i(x, a^*)$. There are solutions for these two cases that take the form (respectively) of the following incentive schemes:

$$\begin{aligned} s_i^+(x) &= \frac{c_i'(a_i^*)}{\mu(\bar{\mathbf{x}}|a^*)\,\lambda_i(\bar{\mathbf{x}}, a^*)} \prod_j \delta(\bar{x}_j - x_j) \\ s_i^-(x) &= \frac{c_i'(a_i^*)}{\mu(\bar{x}_i, 0, a^*)\,\lambda_i(\bar{x}_i, 0, a^*)} \delta(\bar{x}_i - x_i) \prod_{j \neq i} \delta(x_j) \end{aligned}$$

where $\delta(\cdot)$ denotes the Dirac delta function. ■

The following lemma gives a sufficient condition for the first order approach to be valid. The condition will simply guarantee concavity of the problem of each agent given the contract that she faces.

Lemma 26 *If agent i is offered a contract of the form*

$$s_i(x) = \frac{c_i'(a_i^*)}{\mu(\hat{x}, a^*)\,\lambda_i(\hat{x}, a^*)} \prod_j \delta(\hat{x}_j - x_j),$$

then, if for every a_i,

$$\mu_{a_i a_i}\left(\hat{x}|a_i, a_{-i}^*\right) \lambda_i(\hat{x}, a^*) \leq c_i''(a_i)\,\frac{\mu(\hat{x}, a^*)}{c_i'(a_i^*)} \qquad (18)$$

agent i will choose effort a_i^*.

Proof. Given s_i, agent i chooses effort as to maximize

$$E\left[s_i(x)|a_i,a_{-i}^*\right]-c_i(a_i)=\frac{c_i'(a_i^*)}{\mu(\hat{x},a^*)\lambda_i(\hat{x},a^*)}\mu\left(\hat{x}|a_i,a_{-i}^*\right)-c_i(a_i)$$

The first order condition is

$$\frac{c_i'(a_i^*)}{\mu(\hat{x},a^*)\lambda_i(\hat{x},a^*)}\mu_{a_i}\left(\hat{x}|a_i,a_{-i}^*\right)=c_i'(a_i)$$

Noting that

$$\mu_{a_i}(\hat{x}|a)=\mu(\hat{x}|a)\lambda_i(\hat{x},a)$$

it follows straightforwardly that the first order condition is satisfied at $a_i = a_i^*$. Differentiating the F.O.C. wrt a_i, it follows that the agent's objective is concave iff 18 holds. ■

Proof. (of proposition 14)Consider the problem of inducing effort a_1^* for agent 1. The principal sets a prize B that will be awarded to the agent with the highest signal between x_1 and x_2, if the signal is above m. Assume agent 2 exerts effort a_2^*. When exerting effort a, agent 1 will win the contest with probability $\int_m^{\bar{x}} F^2(x|a,a_2^*) f^1(x|a,a_2^*)\,dx$. Therefore she will try to maximize

$$B\int_m^{\bar{x}} F^2(x|a,a_2^*) f^1(x|a,a_2^*)\,dx-c(a)$$

The first order condition for this problem is[10]:

$$\begin{aligned}
c'(a) &= \int_m^{\bar{x}} BF_{a_1}^2(x|a,a_2^*) f^1(x|a,a_2^*)\,dx \\
&\quad+\int_m^{\bar{x}} BF^2(x|a,a_2^*) f_{a_1}^1(x|a,a_2^*)\,dx \\
&= B\int_m^{\bar{x}}\left[\int_0^x f_{a_1}^2(y|a,a_2^*)\,dy\right] f^1(x|a,a_2^*)\,dx \\
&\quad+\int_m^{\bar{x}} BF^2(x|a,a_2^*) f_{a_1}^1(x|a,a_2^*)\,dx
\end{aligned}$$

[10]The notation $a\vee b$ is used to denote $\max(a,b)$. Similarly $a\wedge b$ may be used for $\min(a,b)$.

Changing the order of integrals, and manipulating a little, we get

$$
\begin{aligned}
c'(a) &= B\int_0^{\bar{x}} \left[\int_{m\vee y}^{\bar{x}} f^1(x|a,a_2^*)\,dx\right] f_{a_1}^2(y|a,a_2^*)\,dy \\
&\quad + \int_m^{\bar{x}} BF^2(x|a,a_2^*)\,f_{a_1}^1(x|a,a_2^*)\,dx \\
&= \int_0^{\bar{x}} B\left[1-F^1(m\vee x|a,a_2^*)\right] f_{a_1}^2(x|a,a_2^*)\,dx \\
&\quad + \int_m^{\bar{x}} BF^2(x|a,a_2^*)\,f_{a_1}^1(x|a,a_2^*)\,dx \\
&= \int_m^{\bar{x}} BF^2(x|a,a_2^*)\,\frac{f_{a_1}^1(x|a,a_2^*)}{f^1(x|a,a_2^*)}f^1(x|a,a_2^*)\,dx \\
&\quad - \int_0^{\bar{x}} BF^1(m\vee x|a,a_2^*)\,\frac{f_{a_1}^2(x|a,a_2^*)}{f^2(x|a,a_2^*)}f^2(x|a,a_2^*)\,dx
\end{aligned}
$$

Then, for the first order condition to hold at a^*, it follows that

$$
\begin{aligned}
B &= c'(a_1^*)\left[\int_m^{\bar{x}} F^2(x|a^*)\,\frac{f_{a_1}^1(x|a^*)}{f^1(x|a^*)}f^1(x|a^*)\,dx\right. \\
&\quad \left. - \int_0^{\bar{x}} F^1(m\vee x|a^*)\,f_{a_1}^2(x|a^*)\,dx\right].^{-1}
\end{aligned}
\tag{19}
$$

And note that the cost of the tournament is given by

$$B\left[1-F^1(m|a^*)\,F^2(m|a^*)\right].$$

Consider the case of positive externalities. Noting that $\frac{f_{a_1}^1}{f^1}$ is increasing, and

$$\frac{\int_m^{\bar{x}} F^2(x|a^*)\,f^1(x|a^*)\,dx}{1-F^1(m|a^*)\,F^2(m|a^*)} = \Pr\{x_1 \geq x_2|\max(x_1,x_2)\geq m\} = \frac{1}{2},$$

it follows that

$$\begin{aligned}
& B\left[1-F^{1}\left(m|a^{*}\right)F^{2}\left(m|a^{*}\right)\right] \\
= & c'\left(a^{*}\right)\left[\frac{\int_{m}^{\bar{x}}F^{2}\left(x|a^{*}\right)\frac{f_{a_1}^{1}\left(x|a^{*}\right)}{f^{1}\left(x|a^{*}\right)}f^{1}\left(x|a^{*}\right)dx}{1-F^{1}\left(m|a^{*}\right)F^{2}\left(m|a^{*}\right)}\right. \\
& \left.-\frac{\int_{0}^{\bar{x}}F^{1}\left(m\vee x|a^{*}\right)f_{a_1}^{2}\left(x|a^{*}\right)dx}{1-F^{1}\left(m|a^{*}\right)F^{2}\left(m|a^{*}\right)}\right]^{-1} \\
\geq & c'\left(a^{*}\right)\left[\frac{f_{a_1}^{1}\left(\bar{x}|a^{*}\right)}{f^{1}\left(\bar{x}|a^{*}\right)}\left(\frac{\int_{m}^{\bar{x}}F^{2}\left(x|a^{*}\right)f^{1}\left(x|a^{*}\right)dx}{1-F^{1}\left(m|a^{*}\right)F^{2}\left(m|a^{*}\right)}\right)\right]^{-1} \\
= & 2c'\left(a^{*}\right)\left[\frac{f_{a_1}^{1}\left(\bar{x}|a^{*}\right)}{f^{1}\left(\bar{x}|a^{*}\right)}\right]^{-1}
\end{aligned}$$

The inequality follows because the second integral above is positive (due to positive externalities). Noting that the last expression is equivalent to (7) for two symmetric agents, the first result follows.
For the case of negative externalities, it suffices to show that tournaments can sometimes implement effort a^{*} at a smaller cost than individual bonuses. The next example takes care of this case. ■

Example 27 *An easy way to construct a family of ordered distributions is to take convex combinations of a good distribution $f(\cdot)$ and a 'bad' distribution $g(\cdot)$. For $i\in\{1,2\}$, $j=3-i$, suppose that*

$$f^{i}(x|a)=\frac{1}{2}\left(a_{i}f+\left(1-a_{i}\right)g\right)+\frac{1}{2}\left(a_{j}g+\left(1-a_{j}\right)f\right).$$

That is, one agent's effort has a positive effect over her own signal and a negative effect over the other agent's signal (negative externalities). Let c be positive, increasing and convex with $c'(1)$ sufficiently large. Note that for $a_{i}=a_{j}$,

$$\begin{aligned}
f^{i}(x|a,a) &= f^{j}(x|a,a)=\frac{1}{2}(f+g)\text{, and} \\
f_{a_i}^{i}(x|a) &= \frac{1}{2}(f-g)=-f_{a_j}^{i}(x|a)
\end{aligned}$$

Now, from 7, it follows that the cost of implementing effort $a^{*}=(a,a)$ through individual bonuses is:

$$C^{IPE}(a)=\frac{c'(a)}{2}\frac{f(\bar{x})+g(\bar{x})}{f(\bar{x})-g(\bar{x})}$$

Letting $f(x) = \frac{1}{2} + x$ *and* $g(x) = \frac{3}{2} - x$ *for* $x \in [0,1]$; *it follows that* $f \succeq_{LR} g$, *and* $(f+g)/2$ *is the uniform distribution. Therefore* $C^{IPE}(a) = c'(a)$. *Consider now the cost of implementing* a^* *via a tournament* ($m = 0$)*. The cost is given by 19 and can be expressed in this case as*

$$
\begin{aligned}
C^T(a) &= c'(a) \int_0^1 x[f(x) - g(x)]\,dx \\
&= \frac{c'(a)}{6}
\end{aligned}
$$

Which shows that under negative externalities, tournaments can outperform IPE schemes in providing cheap incentives.

The following two lemmas will be used in the proof of proposition 16. First, it will be convenient to introduce some notation.

Given $x \in \mathbf{R}^n$, and π a permutation of $\{1, ..., n\}$, define $x^\pi = \left(x_{\pi(1)}, ..., x_{\pi(n)}\right)$.

Definition 28 *A compensation scheme* $s : \mathbf{R}^n \to \mathbf{R}^n$ *is said to be symmetric if for all* i, $s_{\pi(i)}(x^\pi) = s_i(x)$.

Lemma 29 *The set of solutions to problem 9 is convex*

Proof. Given a^*, suppose that s and $\hat{s}$ solve 9. For $\alpha \in (0,1)$, it is clear that $\tilde{s} = \alpha s + (1-\alpha)\hat{s}$ attains the same value of the objective function. It is also straight forward to show that it satisfies all the liability constraints. To show that it satisfies the first-order incentive constraints it suffices to note that 4 can be written as:

$$E[s_i(x)\lambda_i(x, a^*)\,|a^*] = c_i'(a_i^*).$$

Hence, if s and $\hat{s}$ satisfy such constraints, $\alpha s + (1-\alpha)\hat{s}$ will also satisfy them. ■

Lemma 30 *For a team of symmetric agents, if there exists a solution to 9, then there exists a symmetric solution. Moreover, any solution can be associated with a particular symmetric solution.*

Proof. Suppose s is a solution to 9. From the symmetry assumption it is clear that for any permutation π, the payment rule s^π defined by:

$$s_i^\pi(x) := s_{\pi(i)}(x^\pi)$$

is also a solution –that is, changing the order of agents has no effect given team symmetry.
Consider all possible permutations of $\{1, ..., n\}$ and denote them as $\{\pi_k\}_{k=1,...,n!}$.
From lemma 29 it follows that $s^* = \frac{1}{n!}\sum_k s^{\pi_k}$ must solve 9. Moreover s^* is symmetric since, given a permutation ρ,

$$\begin{aligned} s^*_{\rho(i)}\left(x^{\rho}\right) &= \frac{1}{n!}\sum_k s^{\pi_k}_{\rho(i)}\left(x^{\rho}\right) \\ &= \frac{1}{n!}\sum_k s_{\pi_k(\rho(i))}\left(x^{\pi_k(\rho)}\right) \\ &= \frac{1}{n!}\sum_k s_{\pi_k(i)}\left(x^{\pi_k}\right) = s^*_i\left(x\right) \end{aligned}$$

where the next to last equality follows from the fact that the π_k's visit all possible permutations, hence it makes no difference to start from ρ than from the original positions. ■

Proof. of proposition 16 I will show that in any solution problem (9), with probability 1, at most one agent receives a payment above the minimum. There is no loss of generality in assuming $L_i = 0$ for all i.

Suppose $\{s^*_i(\cdot)\}_{i=1,...,n}$ is a solution. Define $A_i = \{x : s^*_i > L_i\}$. By way of contradiction, suppose that for some $i \neq j$, $A_{ij} = A_i \cap A_j$ has positive measure: $\Pr(A_{ij}) > 0$. From lemma (30) it suffices to restrict attention to symmetric solutions; assume s^* to be symmetric.
Define m_i, m_j as the marginal incentive power for i and j in A_{ij}, that is:

$$m_i := \int_{A_{ij}} s^*_i(x)\,\mu(x|a^*)\,\lambda_i(x, a^*)\,dx_1...dx_n.$$

By the symmetry assumption and symmetry of the solution, it follows that $m_i = m_j$. Define payment schedules $\hat{s}^{\alpha}$ as follows: for $k \notin \{i, j\}$, $\hat{s}^{\alpha}_k = s^*_k$, and

$$\left(\hat{s}^{\alpha}_i(x), \hat{s}^{\alpha}_j(x)\right) = \begin{cases} \left(s^*_i(x), s^*_j(x)\right) & \text{if} \quad x \notin A_{ij} \\ \alpha\left(s^*_i(x) + s^*_j(x), 0\right) & \text{if} \quad x \in A_{ij}, \lambda_i(x, a^*) \geq \lambda_j(x, a^*) \\ \alpha\left(0, s^*_i(x) + s^*_j(x)\right) & \text{other} \end{cases}$$

Note that $\hat{s}$ is also symmetric. Let B denote the set of realizations x that are more informative about i's effort level than about j's. That is

$$B := \{x : \lambda_i(x, a^*) \geq \lambda_j(x, a^*)\}.$$

Note that, for $\alpha = 1$, the marginal incentive power for i and j on A_{ij} is greater under $\hat{s}^1$ than under s^*. Dropping arguments and writing dx for $dx_1...dx_n$, this can be seen as follows:

$$\begin{aligned}
& \int_{A_{ij}} \left(\hat{s}_i^1 \lambda_i + \hat{s}_j^1 \lambda_j\right) \mu dx \\
= & \int_{A_{ij} \cap B} \left(s_i^* + s_j^*\right) \lambda_i \mu dx + \int_{A_{ij} \cap B^c} \left(s_i^* + s_j^*\right) \lambda_j \mu dx \\
\geq & \int_{A_{ij}} s_i^* \lambda_i \mu dx + \int_{A_{ij}} s_j^* \lambda_j \mu dx = m_i + m_j. \qquad (20)
\end{aligned}$$

The inequality follows from the definition of B. By symmetry, it must be that each agent's incentive power is greater under $\hat{s}$ than under s^*. Therefore, there is some $\alpha \in [0, 1]$ such that

$$\int_{A_{ij}} \left(\hat{s}_i^\alpha \lambda_i + \hat{s}_j^\alpha \lambda_j\right) \mu dx = m_i + m_j.$$

Noting that $\hat{s}^\alpha$ satisfies all the constraints and improves the objective function of (9), it follows that if $\alpha < 1$, s^* could not be a solution which would conclude the proof.
Fixing x_{-i}, note that, from the low positive externalities assumptions, there is at most one value of x_i that satisfies $\lambda_i(x_{-i}, x_i, a^*) = \lambda_j(x_{-i}, x_i, a^*)$. Hence the set $\{x : \lambda_i(x, a^*) = \lambda_j(x, a^*)\}$ has measure zero, and the inequality in (20) is strict –which implies $\alpha < 1$. ■

References

[1] Asch, Beth (1990), "Do Incentives Matter? The Case of Navy Recruiters" *Industrial and Labor Relations Review, V. 43*, 89S-106S.

[2] Athey, Susan, and Armin Schmutzler (2001), "Investment and Market Dominance" *RAND Journal of Economics* 32, 1-26.

[3] Baron, James and David Kreps. (1999) *Strategic Human Resources: Frameworks for General Managers*, John Wiley & Sons.

[4] Bulow, Jeremy; John Genakoplos and Paul Klemperer (1985) "Multimarket Oligopoly:Strategic Substitutes and Complements", *Journal of Political Economy,* V93.3, 388-511.

[5] Cabral, Luis (2002) "Football, Sailing, and R&D: Dynamic Competition with Strategic Choice of Variance and Covariance", *Journal of Economics and Management Strategy ,* forthcoming.

[6] Che, Yeon-Koo and Seung-Weon Yoo (2001), "Optimal Incentives for Teams", *The American Economic Review, V91.3, 525-541.*

[7] Ehrenberg, Ron and Michael Bognanno. (1990) "The Incentive Effects of Tournaments Revisited: Evidence from the European PGA Tour," *Ind. Lab. Relat. Rev.,* 43, p. 74-89

[8] Hart, O. (1995) *Firms Contracts and Financial Structure*, Oxford University Press.

[9] Green, Jerry and Nancy Stokey (1983), "A Comparison of Tournaments and Contracts", *Journal of Political Economy,* V91.3, 349-64.

[10] Holmstrom, Bengt (1979), "Moral Hazard and Observability", *The Bell Journal of Economics,* V10.1, 74-91.

[11] Holmstrom, Bengt (1982), "Moral Hazard in Teams", *The Bell Journal of Economics,* V13.2, 324-40.

[12] Holmstrom, Bengt and Paul Milgrom (1987), "Aggregation and Linearity in the Provision of Intertemporal Incentives", *Econometrica* V55.2, 303-28.

[13] Innes, R. (1990), "Limited Liability and Incentive Contracting with Ex-Ante Action Choices", *Journal of Economic Theory 52, 45-67.*

[14] Innes, R. (1993), "Financial Contracting under Risk Neutrality, Limited Liability and Ex Ante Asymmetric Information", *Economica 60, 27–40.*

[15] Lazear, Edward (1995) *Personnel Economics,* MIT Press, Cambridge Massachusetts.

[16] Lazear, Edward P. and Sherwin Rosen (1981), "Rank-Order Tournament on Optimal Labor Contract" in *Journal of Political Economy.* Vol 89, 841-864.

[17] Levin, Jonathan (2001), "Relational Incentive Contracts", Working paper.

[18] Levin, Jonathan (2002), "Multilateral Contracting and the Employment Relationship", Forthcoming *Quarterly Journal of Economics.*

[19] Mc Afee, Preston and John Mc Millan (1991), "Optimal Contracts for Teams", *International Economic Review,* V32.3, 561-577.

[20] Mirrlees, James (1974), "Notes on Welfare Economics, Information and Uncertainty", in *Essays on Economic Behavior Under Uncertainty*, ed. by M. Balch, D McFadden and S. Wu.

[21] Mirrlees, James (1999), "The Theory of Moral Hazard and Unobservable Behavior:Part I", Review of Economic Studies, V 66, 3-21.

[22] Moldovanu, Benny and Aner Sela (2001), "The Optimal Allocation of Prizes in Contests", *The American Economic Review* V91.3, 542-558.

[23] Mookherjee, Dilip (1984), "Optimal Incentive Schemes with Many Agents", *Review of Economic Studies,* V51.3, 433-446.

[24] Nalebuff, Barry and Joseph Stiglitz (1983*a*), "Prizes and Incentives: Towards a General Theory of Compensation and Competition", *The Bell Journal of Economics,* V14.1, 21-43.

[25] Nalebuff, Barry and Joseph Stiglitz (1983*b*), "Information, Competition, and markets", *The American Economic Review,* V73.2, 278-83.

[26] Oyer, Paul (1998) "Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality", *Quarterly Journal of Economics.* CXIII, 149-185.

[27] Oyer, Paul (2000) "A Theory of Sales Quotas with Limited Liability and rent Sharing", *Journal of labor Economics* V. 18, 405-26.

[28] Prendergast, Canice (1999) "The Provision of Incentives in Firms", *Journal of Economic Literature,* V. 37, 7-63

[29] Shapiro, Carl and Joseph Stiglitz (1984), "Equilibrium Unemployment as a Worker Discipline Device", *The American Economic Review,* V74.3, 433-44.

[30] Shavell, Steven (1979) "Risk Sharing and Incentives in the Principal and Agent Relationship" *The Bell Journal of Economics*, Vol. 10, No. 1., pp. 55-73.

[31] Zenger, Todd (1992), "Why do Employers Only Reward Extreme Performance? Examining the Relationship Among Performance, pay, and Turnover", *Administrative Science Quarterly,* V37.2, 198-219.