Correlation and Relative Performance Evaluation

Pierre Fleckinger*

March 2008

Abstract

This paper reexamines the issue of relative versus joint incentive schemes in a multiagent moral hazard framework. It allows to contrast the inference dimension and the insurance property of relative performance evaluation. Importantly, the widespread idea that the principal should use all the more competitive schemes that the equilibrium outcomes are more correlated is shown not to be robust. When correlation varies with the efforts chosen, more *equilibrium* correlation can make joint performance evaluation more likely to be optimal, because a pair of good performances can become a relatively better signal that both agents work hard than a pair of asymmetric performances. With risk-averse agents, that informational effect has to be traded off against the agents' insurance concerns. As a result, the optimal incentive scheme is sometimes mixed, which can be interpreted in a firm context as the use of aggregate profit sharing in combination with selective firing or promotion.

Keywords: Multi-agent moral hazard; Correlation; Uncertainty; Relative vs Joint Performance Evaluation.
JEL Classification: D86.

*Columbia Business School and Ecole Polytechnique. Email: pierre.fleckinger@polytechnique.edu.
For comments and suggestions, I am grateful to Yeon-Koo Che, Philippe Février, François Larmande, Trond Olsen, Jean-Pierre Ponssard, Wilfried Sand-Zantman, Jean-Marc Tallon and especially to Patrick Bolton and Bernard Salanié. I also thank participants to EARIE 2007 (Valencia), the *Tournaments, Contests and Relative Performance Evaluation* Conference at NCSU and seminar participants at CREST-LEI and Columbia University. Partial support from the Ecole Polytechnique Chair in Business Economics is gratefully acknowledged.

1 Introduction

When the performances of many agents are subject to correlated shocks, agency theory traditionally indicates that competition between them allows to reduce incentive costs. This paper offers a reconsideration of this view by giving a closer look at the informational structure of multiagent moral hazard settings. In particular, the analysis unveils an informational effect of correlation that runs counter to the traditional perspective. The common view says the following: if performances are positively correlated, a good result of one agents indicates a favorable environment, and therefore one agent's compensation should depend negatively on the other's performance in order not to reward luck. The alternative view says in addition that if performances are more correlated when agents exert more effort, a good performance by both agents becomes a better signal of high efforts than a pair of asymmetric results. This second effect may more than offset the first one, and make joint incentives optimal even under positively correlated performances. The analysis shows that the common view is indeed misleading–or at least incomplete–in that it does not allow for the correlation to vary with the choice of action.

As an illustration (developed later as an example), consider the case of stock-options in startups. Agents' performances are positively correlated because of the underlying technology they are exploiting. A straightforward application of the traditional idea would therefore recommend to use relative performance evaluation to filter the common underlying uncertainty by comparison. The widespread use of stock-options–a joint incentive scheme–in that context seems therefore at odd with the theory. The results of this paper helps rationalize this in the case of risk-averse agents subject to limited liability. A simple intuition for this case might be grasped in an analogy with information gathering. If two agents are asked to report pieces of information on the same underlying variable, and the precision of their signal is increasing with effort, rewarding them when their signals coincide, i.e. jointly, is better than rewarding them when they differ, i.e. relatively. Clearly in that case more effort means more correlation, and this is crucial in designing the incentive schemes. It is shown that similar insights apply to the usual moral hazard model.

With multiple agents, the literature has since the beginning emphasized the role of relative performance evaluation, prominently the early tournament literature, e.g. Lazear and Rosen (1981), Green and Stokey (1983) and Nalebuff and Stiglitz (1983). Those authors have shown in particular that competition between agents is all the more valuable that the common risk associated with individual production increases.[1] This is rooted in the multiagent application of Holmström (1979) sufficient statistics result, by Holmström (1982) and Mookherjee (1984). However, this key result only asserts that the reward for one agent should depend on the performance of the other agent, but by no means commands that the type of evaluation should be relative. It is often mistakenly taken as an argument for provision of competitive incentives. From a broad theoretical point of view, results are not that clear-cut. On the one hand works on general stochastic structures in multiagent problems–e.g. Mookherjee (1984); Ma (1988); Brusco (1997) and d'Aspremont and Gérard-Varet (1998) for the case of partnerships–can not provide answer on the desirability of competition in teams, and on the other hand, works focused on that topic consider rather specific informational structures and technologies (e.g. Maskin et al., 2000; Che and Yoo, 2001; Luporini, 2006) and/or restricted contract forms (e.g. Holmström and Milgrom, 1990; Ramakrishnan and Takor, 1991; Itoh, 1992). The setting analyzed in the following contributes to filling this gap by providing a framework that generalizes the informational structure in two directions,[2] while still allowing for a full characterization of the optimal incentive scheme.

Following Che and Yoo (2001), we will use the terminology relative performance evaluation (RPE) and joint performance evaluation (JPE) to distinguish the two usual ways of paying the agents.[3] The core issue is to characterize situations in which one or the other kind of scheme is the optimal one. We consider both the cases of risk-neutral agents subject to limited liability, and the case of risk-averse agents under a standard participation constraint. Those two cases constitute in fact the two sides of the same coin. This allows to decompose the problem between the two main effects at work: the informational ef-

[1] This is also the case with correlated private information. The idea is that competition allows to cross-check messages and reduce rents, see Demski and Sappington (1984) for an early contribution.

[2] For example, Maskin et al. (2000) restrict attention to noise independent of effort, both regarding variances and correlation. More importantly–given the widespread use–the linear-exponential-normal model has a uniform correlation over effort pairs. In Mookherjee (1984) and Ramakrishnan and Takor (1991), the papers closest to our setting regarding the information structure, the level of correlation is in fact the same irrespective of the action chosen, see Mookherjee (1984, pp. 441-442) and Ramakrishnan and Takor (1991, pp. 259-260, in particular the beginning of section 4.2)

[3] We abstract from (potentially beneficial) cooperative agreements between the agents. Side-transfers and better observation among the agents may by themselves be reasons for cooperative schemes, see for example Holmström and Milgrom (1990) and Itoh (1992). See also Baliga and Sjöström (1998) for a model of sequential efforts in a limited liability framework close to the one developed here. Their focus is however on collusion, a topic not treated here.

fect, by which the principal gets information on the technology and infers effort, and the insurance effect associated with relative performance evaluation.

After setting the model and giving its basic properties, we study in more details an information structure that we refer to as the case of technological uncertainty. It corresponds to case in which the probabilities of obtaining a given result are commonly but imperfectly known. In that setting, correlation comes from ex-ante beliefs on the quality of the technology and create correlation between the outcomes of the agents. As is shown, this makes the information structure more flexible in terms of correlation and variance than the classic approach with perfectly known technology and ad hoc correlation of noisy results. One should also underline the link with some models of ambiguity. As in traditional in the agency literature, the model stays within the comfortable framework of subjective expected utility and Bayesian probabilities. That is we use the Anscombe and Aumann (1963) framework, and therefore preserve the reduction of compound lotteries property and keep additivity of all probability (i.e. we do not use Gilboa and Schmeidler (1989) or Ghirardato et al. (2004) constructions to introduce uncertainty aversion). But the multiagent nature of the framework is itself the source of a Bayesian ambiguity effect in the problem. In that respect, the closest paper is Halevy and Feltkamp (2005). Essentially, they extend Ellsberg's famous thought experiment by one draw: bets are on two successive draws from the same urn. Their results is that under such circumstances, a purely Bayesian decision-maker will be *uncertainty*-averse as soon as he is *risk*-averse. Roughly speaking,[4] the present model has in common with that paper the "double draw effect". An interesting aspect is that even under risk-neutrality of the agents, the principal can use uncertainty, and he is in fact better off with ambiguity than without.

In the next section, we set up the model and state the main definitions. Then in the third section we derive the optimal contract under risk-neutrality and limited liability and apply it to various examples. The fourth section deals with risk-aversion and insurance. The last section concludes. Omitted proof are relegated to the appendix.

[4]This is a rough parallel in that it abstracts from incentive concerns.

2 Model under risk-neutrality and preliminary analysis

2.1 Basics

We consider a setting of moral hazard in which two agents, called 1 and 2, work on two different projects. The principal and the agents share the same beliefs on the technology. All players are risk-neutral in the present and the following section, and we relax this assumption in the fourth section by considering risk-averse agents. Following Itoh (1991), Ramakrishnan and Takor (1991) and Che and Yoo (2001), the outcome of each project is either a success or a failure, worth respectively S and F to the principal,[5] and these outcomes are contractible. We denote by $\boldsymbol{R}$ a generic result pair, taking value in $\{SS, SF, FS, FF\}$. Each agent privately chooses whether he exerts effort or not: agent 1 chooses $e \in \{0,1\}$ and agent 2 chooses $f \in \{0,1\}$. Those actions are never observed by the other players. We assume without loss of generality that the costs functions are $c(e) = c.e$ and $k(f) = k.f$ for agent 1 and 2 respectively. The probability of obtaining outcome $\boldsymbol{R}$ conditional on effort pair (e, f) is $Prob[\boldsymbol{R}|e, f]$. Finally, we consider only projects for which it is worth inducing the agents to work, which amounts to assume that $S - F$ is sufficiently high compared to c. This also renders the assumption of identical payoffs and costs innocuous.

2.2 Incentive Schemes

The performances of the agents are generically related, so that their wages should be tied. The incentive scheme (or wage profile) for agent 1 is thus a collection

$$\boldsymbol{w} = \{w_{SS}, w_{SF}, w_{FS}, w_{FF}\}$$

that represents the wage he receives contingent on his result–the first index–and on the second agent's result–the second index. The wage scheme of agent 2 is denoted by $\boldsymbol{x}$ and follows the same conventions. In practice, the problem is symmetric and separable and we mainly focus on the case of agent 1 only.[6] Given an outcome-contingent wage scheme $\boldsymbol{w}$ and a pair of efforts (e, f), the expected payoffs are:

[5] The results do not depend on this symmetry, it is assumed only for clarity.

[6] We show in the first lemma that the problem being separable for the principal, we can in full generality focus on a single agent.

$$U_1(\boldsymbol{w}|e,f) = \mathbb{E}_{\boldsymbol{R}}\left[w_{\boldsymbol{R}}|e,f\right] - c(e)$$
$$U_2(\boldsymbol{x}|e,f) = \mathbb{E}_{\boldsymbol{R}}\left[x_{\boldsymbol{R}}|e,f\right] - k(f)$$

From each agent's point of view, there are overall three stochastic sources in that payment: first, there is imperfect knowledge on the technology–a specificity of this paper, second, the result of his effort is non-deterministic–as is standard in moral hazard settings, and third his remuneration also depends on the other agent's stochastic performance–a feature of the two-agent setting. The expectation operator pertains to those three random elements. Since there is a finite number of outcomes, we can write the expected wage as a weighted average, where the weight on $w_{\boldsymbol{R}}$ is $Prob(\boldsymbol{R}|e,f)$.

A central question of the analysis is: When will the principal use competition as an incentive device, and when will he prefer to induce cooperative behavior between the agents? To give a precise content to this question, we borrow from Che and Yoo (2001) the typology for the incentive systems.

Definition 1 *(Standard incentive schemes)*
*An incentive scheme exhibits **Relative Performance Evaluation** (RPE) when:*

$$(w_{SF}, w_{FF}) > (w_{SS}, w_{FS})$$

*An incentive scheme exhibits **Joint Performance Evaluation** (JPE) when:*

$$(w_{SS}, w_{FS}) > (w_{SF}, w_{FF})$$

*An incentive scheme exhibits **Independent Performance Evaluation** when:*

$$(w_{SS}, w_{FS}) = (w_{SF}, w_{FF})$$

The inequalities represent component-wise comparison with at least one strict inequality. With RPE, an agent is better off when the other fails, while it is the converse with JPE. Therefore, RPE is competitive while JPE gives collaborative incentives.[7] Note that these three types of scheme do not exhaust the possible ordering of wages, and indeed additional configurations appear in the fourth section.

[7] Strictly speaking, there can be no collaboration, since an agent can not influence the other's result in this model beyond his single effort choice. Accounting for wider possibilities would require to enrich the action space to account for sabotage (Lazear, 1989) or help (Itoh, 1991). Note also that JPE and RPE raise different collusion issues, see Brusco (1997) for a general approach to collusion.

2.3 Implementation and toolbox

As already mentioned, it is assumed that the principal wants both agent to exert effort, and we limit ourselves to minimizing the cost of implementing $(1,1)$ as a Nash equilibrium (not necessarily a unique one). The incentive constraints for each agent provided the other exerts effort write:

$$U_1(\boldsymbol{w}|1,1) \geq U_1(\boldsymbol{w}|0,1) \tag{1}$$

$$U_2(\boldsymbol{x}|1,1) \geq U_2(\boldsymbol{x}|1,0) \tag{2}$$

In addition, we assume that the agents are subject to limited liability, so that the following constraints hold:

$$\boldsymbol{w}, \boldsymbol{x} \geq 0 \tag{3}$$

Since he wants to implement the effort pair $(1,1)$, the principal's objective is to minimize the sum of transfers under the previous constraints:

$$\min_{\boldsymbol{w},\boldsymbol{x}} \mathbb{E}_{\boldsymbol{R}} \left[w_{\boldsymbol{R}} + x_{\boldsymbol{R}} | 1,1 \right]$$

$$\text{subect to } (1), (2), (3)$$

It is however possible to deal with a simplified problem as the next observation indicates.

Lemma 1 *The program of the principal is separable in two independent optimization prolems, one for each agent.*

This result is useful for simplifying the exposition. The separability follows from two facts: First, since the principal is risk neutral, the objective function is linear in wages, and second, the constraints feature the wages of only one agent at a time, and they can thus be divided into two independent sets. In the following, we consequently restrict attention to the program pertaining to agent 1. In other words, we consider only one side of the problem, but the other can be dealt with in the exact same way. Since for the remaining of the analysis we stick to agent 1's point of view, the following definitions are implicitly given for agent 1 only. Those definitions will help grasping some economic intuition during the resolution. They are simply the specialized definitions of the likelihood ratio for this model.

Definition 2 *For any pair of results $\boldsymbol{R}$, the **effort informativeness** of $\boldsymbol{R}$ is:*

$$h(\boldsymbol{R}) = \frac{Prob(\boldsymbol{R}|1,1)}{Prob(\boldsymbol{R}|0,1)}$$

To clarify why we call informativeness this ratio, it is illustrative to consider one of them, $h(SS)$. It is the likelihood ratio between effort and shirking for the first agent upon observing two successes, conditional on the other agent exerting effort. The higher $h(SS)$ is, the more likely it is upon observing two successes that the agent has exerted effort and not shirked. The different possible results carry more or less information about the choice of effort, and h is a measure of that information. Note that in this definition we only consider cases in which agent 2 exerts effort, since in the end we are concerned with the implementation of two efforts.

Definition 3 *For any pair of results $\boldsymbol{R}$, the **incentive efficiency** of $w_{\boldsymbol{R}}$ is:*

$$I(\boldsymbol{R}) = 1 - \frac{1}{h(\boldsymbol{R})} = \frac{Prob(\boldsymbol{R}|1,1) - Prob(\boldsymbol{R}|0,1)}{Prob(\boldsymbol{R}|1,1)}$$

The right-hand side is the ratio between the coefficient of the wage $w_{\boldsymbol{R}}$ in the incentive constraint and the probability of paying this wage. This is therefore the (constant) ratio of marginal incentive and marginal costs for that wage, which explains the notion of incentive efficiency. Note that it is at most 1, in which case the wage is fully effective, since the result $\boldsymbol{R}$ then indicates with certainty that the agent has exerted effort.

Two remarks are in order about these two definitions. First, in a setting with one single agent and binary outcomes, these two concepts are trivial, but they make sense when the outcome space in richer, as with multiple agents or additional signals (see Laffont and Martimort, 2002, pp. 167-172). Second, note that a result $\boldsymbol{R}$ is more informative than a result $\boldsymbol{R}'$ if and only if the associated wage $w_{\boldsymbol{R}}$ is more incentive efficient than $w_{\boldsymbol{R}'}$. In other words, it is equivalent to reason in terms of how to best infer the action or in terms of how to spread optimally the incentive weight. The next result will be useful in characterizing the optimal incentive scheme.

Lemma 2 *Under risk-neutrality and limited liability, an optimal incentive scheme entails positive wages only for the result(s) with the highest incentive efficiency.*

The lemma expresses the intuitive idea that the incentive weight should be put on the outcomes that are most efficient at inducing effort. In fact, since the principal's objective is linear, the result is even more extreme, and all the weight will generically be put on one single wage.

We are now in position of grasping a generic flavor of the situations under which the optimal scheme exhibits RPE or JPE. Another definition is needed, once again limiting us to the point of view of agent 1.

Definition 4 *Effort e is a **strong complement** (resp. **substitute**) to effort f if:*

$$\frac{Prob(.S|1,1)}{Prob(.F|1,1)} \geq (\leq) \frac{Prob(.S|0,1)}{Prob(.F|0,1)}$$

That is, when agent 1 exerts effort, this increases the likelihood of a success relative to a failure for agent 2 (for any result of agent 1). Note that this relationship is not symmetric, since it may well be the case that effort f is detrimental to the success of the first project. In any case, we have the following result.

Proposition 1 *The optimal incentive scheme of agent 1 exhibits JPE (resp. RPE) if and only if effort e is a strong complement (resp. substitute) to effort f.*

Keeping as a reference this generic insight that collective schemes are optimal under (some notion of) complementarity,[8] and relative performance evaluation optimal under substitutability, we investigate now the information structure in more details.

3 Main result under risk-neutrality

3.1 A heuristic approach

We first give an insight of the results in a simple way. No restrictions where yet imposed on how efforts interact in the production process. From now on, we will focus exclusively

[8] It is worthwhile observing that Aggarwal and Samwick (1999) obtain a related result in the case of competing structures, each with one principal and one agent. When firms compete à la Cournot, the actions of the agents–quantity choices–are strategic substitutes, and RPE is desirable for the principals, while JPE is optimal under Bertrand competition, under which agents' actions are strategic complements. In fact, the present model could also be applied to competing structures, to the extent that as each principal would want his agent to exert effort. Indeed, in such a case, both agents will exert effort in equilibrium and therefore the incentive constraint in each competing structure is the same as that for each agent in an integrated structure. We do not elaborate here on this issue, but predictions can be derived in the field of top executives compensation.

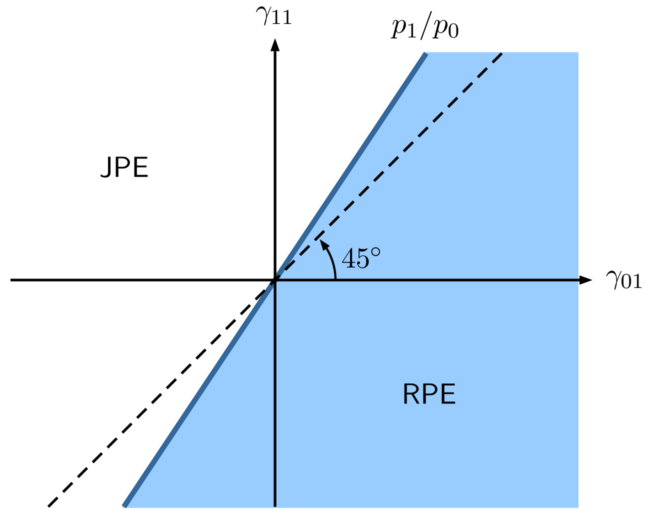

Figure 1: Optimal scheme in the covariance space

on the informational dimension of the problem, by assuming away 'technological' interaction, i.e. the effort of one agent does not influence the result of the other agent. This implies that we have in the following:

$$Prob(R_1 = S|ef) = Prob(R_1 = S|e) \equiv p_e$$

and we denote similarly q_f the probability of a success of agent 2 conditional on exerting effort f. In general, we can thus describe the distribution on the outcome pair by the table:

	S	F
S	$p_e q_f + \gamma_{ef}$	$p_e(1-q_f) - \gamma_{ef}$
F	$(1-p_e)q_f - \gamma_{ef}$	$(1-p_e)(1-q_f) + \gamma_{ef}$

Where γ_{ef} is the covariance of the outcomes when the effort pair is (e, f). Anticipating a bit, and assuming that the relevant likelihood rations are $h(SS)$ and $h(FS)$, it is straightforward to obtain that relative performance is optimal when:

$$h(SF) > h(SS) \Leftrightarrow p_0\gamma_{11} \leq p_1\gamma_{01}$$

This is represented in the next picture.

To understand the conventional wisdom, one can locate it on the figure. The usual results are attached to a situation around the 45 degree line. On the diagonal, precisely, the optimal is scheme is relative performance evaluation if and only if the (constant) covariance is positive. In fact, a constant correlation corresponds to the following case. First notice that:

$$var(R_1|e) = p_e(1-p_e) \quad \text{and} \quad var(R_2|f) = q_f(1-q_f)$$

therefore, for constant correlation level r, one can write:

$$\gamma_{ef} = r\sqrt{var(R_1|e)var(R_2|f)}$$

and, assuming $r > 0$ we have:

$$\frac{\gamma_{11}}{\gamma_{01}} = \sqrt{\frac{p_1(1-p_1)}{p_0(1-p_0)}} < \frac{p_1}{p_0}$$

which says that RPE is optimal for any positive r. (It is straightforward to treat the symmetric case of optimal JPE with negative correlation). In other words, constant correlation pins down the covariances, so that they are in the zone between the diagonal and the line with slope p_1/p_0.

3.2 Technological uncertainty

We introduce now an information structure displaying technological uncertainty. We assume that the probabilities of success conditional on effort are not known with certainty.[9] All players hold the same beliefs on that probabilities. For agent 1, the success probability conditional on effort e is the random variable $\tilde{p}_e$, and it is $\tilde{q}_f$ for agent 2. In the following we use the notations:

$$p_e = \mathbb{E}[\tilde{p}_e], \qquad \sigma_e^2 = var(\tilde{p}_e), \qquad q_f = \mathbb{E}[\tilde{q}_f], \qquad \tau_f^2 = var(\tilde{q}_f)$$

We allow any correlation between the results in case of high effort and low effort, but the following correlation parameters are central:

$$\rho_{e,f} = \frac{cov(\tilde{p}_e, \tilde{q}_f)}{\sigma_e \tau_f}$$

[9] The reader will notice the link with the classical two-armed bandit problem. Two-armed bandits are classically equipped with one safe and one risky arm, see for example Bolton and Harris (1999). DeGroot (1970, pp. 399-405) contains a model of two-armed-bandit with dependent arms which is the closest setting up to our knowledge. However it consists of an example in which the two Bernoulli parameters have two points supports.

This is the most general form of imperfect knowledge one can introduce in the present setting. It is related to recent ideas in modelling ambiguity (see Klibanoff et al., 2005). Note that the correlation coefficients pertain to the beliefs on the distribution of results but, as is shown below, it also represents the correlation between outcomes .

How ever intuitive lemma 2 can be, it does not give any indication on which wage will be positive. In fact, it still leaves the door open to much less intuitive indications regarding optimal schemes. The next example illustrates this point.

Example 1 *Extreme innovation.*

Consider symmetric agents using identical technologies. Assume that an old well-known technology yields at no cost a success with fixed probability $p_0 (= q_0)$. In turn, the new technology might be either a perfect fit, with probability p_1, or be completely useless, with probability $1 - p_1$. A perfect fit yields a success with probability 1, while a useless technology yields a failure for sure. Implementing the new technology requires to incur a learning cost c. Here $\tilde{p}_1 (= \tilde{q}_1)$ is a binomial distribution with parameter p_1–and consequently variance $\sigma_1^2 = p_1(1 - p_1) > 0$. Therefore, we have $h(SS) = \frac{p_1^2 + \sigma_1^2}{p_1 p_0} = \frac{p_1^2 + p_1(1-p_1)}{p_1 p_0} = \frac{1}{p_0}$, $h(SF) = h(FS) = 0$ and $h(FF) = \frac{(1-p_1)^2 + \sigma_1^2}{(1-p_1)(1-p_0)} = \frac{1}{1-p_0}$. Thus if $p_0 > \frac{1}{2}$, the highest likelihood ratio is that of a double failure: agents should be compensated only in that state.

There may thus be counter-intuitive situations in which agents are rewarded only upon obtaining two failures, such as that described in example 1. Note however that in example 1 the Nash equilibrium that the principal would like to implement may be Pareto-dominated from the agents point of view. Indeed, for small enough p_0 relative to p_1, they are better off in the shirk-shirk equilibrium (for any w_{FF}). Tacit coordination of the agents on their Pareto-optimal equilibrium[10] could be by itself a reason not to use a scheme with $w_{FF} > 0$, since this would encourage the worst behavior from the principal's point of view. We will simply rule out such situations by the following assumption on the technology, and we will not impose here unique implementation.

Assumption 1 *(Effective Effort)* $\quad Prob(\tilde{p}_1 \geq \tilde{p}_0) = 1$

[10] The topic of unique and/or undominated implementation is of interest in itself. It has been thoroughly studied since the early contributions of Ma (1988) and Ma et al. (1988). For the closest settings, see in particular Arya and Glover (1995) and Arya et al. (1997).

This quite natural assumption says that any draw for the technology in case of effort is better than any possible draw in absence of effort. Therefore, exerting effort always (weakly) increases the probability of success. This does not mean that effort is always efficient, since efficiency considerations should also take into account the cost c. Example 1 is ruled out because in some states success could be less likely when the agents exerted effort. The assumption is stronger than needed for our results, but it has the advantage of being perfectly transparent. A useful consequence of assumption 1 is the following.

Lemma 3 *Under assumption 1, an optimal scheme entails* $w_{FF} = w_{FS} = 0$.

In fact, for any result of agent 2, the assumption implies that a failure of agent 1 is always ranked below a success in terms of incentive efficiency.

3.3 Main result

We are now in position to fully characterize the optimal incentive scheme. This is done in the next proposition, using the definition:

Proposition 2 *Under assumption 1, the optimal wage profile is:*

if $\rho_{11}\frac{\sigma_1}{p_1} < \rho_{01}\frac{\sigma_0}{p_0}$, *a RPE scheme with*

$$w_{SF} = \frac{c}{(1-q_1)(p_1-p_0) - \tau_1(\rho_{11}\sigma_1 - \rho_{01}\sigma_0)}, \quad w_{SS} = w_{FS} = w_{FF} = 0$$

if $\rho_{11}\frac{\sigma_1}{p_1} > \rho_{01}\frac{\sigma_0}{p_0}$, *a JPE scheme with*

$$w_{SS} = \frac{c}{q_1(p_1-p_0) + \tau_1(\rho_{11}\sigma_1 - \rho_{01}\sigma_0)}, \quad w_{SF} = w_{FS} = w_{FF} = 0$$

if $\rho_{11}\frac{\sigma_1}{p_1} = \rho_{01}\frac{\sigma_0}{p_0}$, *any scheme (possibly IPE) with*

$$(q_1 + \frac{\tau_1}{p_0})w_{SS} + (1 - q_1 - \frac{\tau_1}{p_0})w_{SF} = \frac{c}{p_1-p_0}, \quad w_{FS} = w_{FF} = 0$$

The criterion for relative vs joint performance evaluation, though it looks very simple, has a rich economic content. On the technical side, it is completely generic in that it does not depend on any assumption on the shape of the underlying distributions,[11] nor on

[11] In the multiagent models mentioned in the introduction, almost all distributions are either multivariate normal representing additive noise or two-point distributions.

whether they are discrete or continuous, and so on. Also, it depends only on the properties of the distributions, so that it is purely informational, while proposition 1 included the technological dimension. We can however relate this result to the link between effort complementarity and JPE demonstrated in the first proposition. We have on the one hand:

$$\frac{Prob(SS|1,1)}{Prob(SF|1,1)} = \frac{p_1 q_1 + \rho_{11}\sigma_1\tau_1}{p_1(1-q_1) - \rho_{11}\sigma_1\tau_1}$$

which is increasing in ρ_{11} and on the other hand:

$$\frac{Prob(SS|0,1)}{Prob(SF|0,1)} = \frac{p_0 q_1 + \rho_{01}\sigma_0\tau_1}{p_0(1-q_1) - \rho_{01}\sigma_0\tau_1}$$

which is increasing in ρ_{01}. Following definition 4, the effort of the first agent is all the more complementary (for successes) to that of the second agent that ρ_{11} is relatively higher than ρ_{01}. According to proposition 1, such complementarity calls for JPE. In that sense, proposition 2 emphasizes that correlation creates partial *informational* complementarities.

Proposition 2 makes an important connection between two dimensions: First, the correlation conditional on effort and second the effect of effort on the variability of the result. In the case of positive correlation, the criterion for RPE can be rewritten as follows:

$$\frac{\rho_{11}}{\rho_{01}} > \frac{p_1}{\sigma_1}\frac{\sigma_0}{p_0}$$

The first dimension is a pure multiagent effect, while the second is a pure single agent effect. The coefficient of variation $\frac{\sigma_e}{p_e}$ is a measure[12] of how noisy the success signal is as a function of effort e. If the term on the right hand side is smaller than 1, effort increases this noise,[13] while it decreases noise if it is higher than 1. Regarding correlations, if the ratio $\frac{\rho_{11}}{\rho_{01}}$ is higher than 1, the results of the agents are more correlated when they choose the same actions than when they choose different actions. This ratio is new in the multiagent analysis since previous papers consider uniform correlation.[14] What matters in the choice

[12]Many related measures are defined in financial analysis, such as the Sharpe and the information ratios. The inverse of the coefficient of variation is referred to as the 'Sharpe ratio' in portfolio analysis. Portfolios with a smaller Sharpe ratios are considered riskier, that is, noisier in the informational interpretation of the model.

[13]Whether exerting more effort increases noise or not and what are the consequences for the career concern model is a point discussed in Dewatripont et al. (1999).

[14]Ramakrishnan and Takor (1991) insist on the role of *conditional* correlation, but as they mention p. 260, the agents take the value of the correlation as exogenous in their model. It seems reasonable to assume that sophisticated agents take into account the fact that the correlation varies with their own choice of action.

between RPE and JPE is the relative quality of information between the two possible actions for one agent, and how the correlation between results varies with the different actions.

Remark 1 *It is possible to reformulate the criterion as follows. First, define the following variable change:*

$$\tilde{p}_e = p_e(1+\varepsilon_e) \text{ and } \tilde{q}_f = q_f(1+\eta_f)$$

where the ε's and the η's are random variables with zero means. Then the optimal scheme will be JPE if and only if:

$$cov(\varepsilon_1, \eta_1) > cov(\varepsilon_0, \eta_1).$$

However, this expression does not allow to disentangle between the effect of the dependence of variance on the level of effort and the effect of the dependence of correlation on the pair of efforts.

We finish this discussion by two corollaries to proposition 2.

Corollary 1 *An increase in the equilibrium correlation of the results favors JPE.*

Proof. The equilibrium covariance of the two results is $cov(\tilde{p}_1 S + (1-\tilde{p}_1)F, \tilde{q}_1 S + (1-\tilde{q}_1)F) = (S-F)^2 \rho_{11}\sigma_1\tau_1$, while the variances of the results are $var(\tilde{p}_1 S + (1-\tilde{p}_1)F) = (S-F)^2\sigma_1^2$ and $var(\tilde{q}_1 S + (1-\tilde{q}_1)F) = (S-F)^2\tau_1^2$. The equilibrium correlation is thus exactly ρ_{11}, and an increase in ρ_{11} increases the desirability of JPE, according to proposition 2. ■

This observation runs counter to usual results. This is the most striking consequence of the informational complementarity effect. Other models focus on the idea that good performances indicate a favorable environment and that this favorable noise should be filtered by RPE, while a previously ignored effect is here at work: A good result of the other agent might also be a good signal of effort under high equilibrium correlation.

Corollary 2 *The principal always benefits from uncertainty on the technology.*

Proof. The expected gains of the principal do not depend on uncertainty. While with perfect knowledge of the technology, the principal would use independent schemes, with imperfect knowledge, he can still use a pair of independent contract (IPE) but it is not optimal. Therefore he earns more with RPE or JPE and thus benefits from the uncertainty.

■

It has already been remarked a number of times in the literature that correlation helps reducing rents, smoothes information revelation and so on. Here, the same positive conclusion can be drawn with respect to uncertainty in general, provided all the players are risk-neutral.

3.4 Examples

To illustrate the main result, we briefly apply it to usual forms of uncertainty that have been considered in the literature. The optimal scheme is RPE in the first two examples, while it is IPE or JPE in the last two ones.

Example 2 *The additive model.*

The most classical way of introducing technological uncertainty amounts in our discrete model to the following form of uncertainty:

$$\tilde{p}_e = p_e + \varepsilon$$

$$\tilde{q}_f = q_f + \eta$$

where ε and η are random variables with zero means, variances σ^2 and correlation coefficient ρ. What matters here is that the noise is additively separable from the influence of the action. Note that the variance of the probability of success depends only on ε, which implies $\sigma_0 = \sigma_1 = \tau_0 = \tau_1 = \sigma$. All pairs $(\tilde{p}_e, \tilde{q}_f)$ have the same correlation ρ. Also, $\frac{\sigma_0}{p_0}\frac{p_1}{\sigma_1} = \frac{p_1}{p_0} > 1$. Therefore, according to proposition 2, RPE is always optimal with additive uncertainty. That formulation of additive uncertainty parallels that of Lazear and Rosen (1981) and Nalebuff and Stiglitz (1983). We have shown that this setting favors competition between agents, even abstracting from risk-sharing concerns.

Example 3 *Effort is sometimes irrelevant.*

Che and Yoo (2001) use an original information structure. They assume that with some probability ν, the technology is such that a project is a success regardless of effort,[15] and

[15] It is unimportant that the probabilities are exactly 1, what matters is that they are the same in some state of the world, making effort irrelevant in that state.

with probability $(1-\nu)$, the technology is such that the outcome depends on the effort. Let us denote the probability of success in that case by r_e. Overall, this corresponds to the situation:

$$\{\tilde{p}_0, \tilde{p}_1\} = \{\tilde{q}_0, \tilde{q}_1\} = \begin{cases} \{1, 1\} & \text{with probability } \nu \\ \{k_0, k_1\} & \text{with probability } (1-\nu) \end{cases}$$

The relationships with our notations are simply:

$$k_e = \frac{p_e - \nu}{1-\nu}, \qquad \sigma_e^2 = \frac{\nu}{1-\nu}(1-p_e)^2 \qquad \text{and} \qquad \rho_{ef} = 1$$

Therefore, since we have here $\frac{p_1(1-p_0)^2}{p_0(1-p_1)^2} > 1$, this setting generates RPE as optimal incentive scheme.[16]

Example 4 *The multiplicative model.*

Consider the following setup where the probability of success of an agent is given by:

$$\tilde{p}_e = \varepsilon p_e$$

$$\tilde{q}_f = \eta q_f$$

where ε and η are random variable with means 1, variances σ^2 and correlation ρ. This functional form of uncertainty is used as an example in Holmström (1982) and appears also in career concerns model (e.g. Dewatripont et al., 1999). Note that all pairs $(\tilde{p}_e, \tilde{q}_f)$ have the same correlation ρ, so that $\frac{\rho_{11}}{\rho_{01}} = 1$. In addition we have $\sigma_0 = \tau_0 = r_0\sigma_\varepsilon$ and $\sigma_1 = \tau_1 = r_1\sigma_\varepsilon$. Thus IPE is an optimal scheme in that case (but not the unique one as seen in the proposition). Note that this is true for any level of equilibrium correlation ρ.

Example 5 *Stock-options in startups.*

As a last application, consider the problem of inducing innovation in a new technological venture. A status quo solution consists in using an old, known technology, which generates a success with fixed probabilities p_0 and q_0 (so that $\sigma_0 = \tau_0 = 0$). In turn, the agents

[16] To be clear, the contribution of Che and Yoo (2001) is precisely to show that, while in this static setting RPE is optimal, JPE becomes optimal in the infinitely repeated version of the problem. One contribution of this paper is to give a definite answer to their footnote 16, p. 530-531 regarding how the form of the common shocks affects the choice RPE vs JPE in the static setting.

can innovate, at a cost c, in order to use a new, imperfectly known technology, characterized by the random probability of success $\tilde{p}_1 = \tilde{q}_1$. In that case, $\frac{\rho_{11}}{\rho_{01}}$ tends to infinity, thus proposition 2 indicates that the optimal way of inducing technological change is to use a JPE scheme. There are admittedly other reasons that explain the use of stock options in startups, but at least this example shows that it may even be profitable for a principal (say, a venture capitalist) to incentivize collectively the member of the startup for a purely informational reason.

4 Risk-averse agents and mixed schemes

Now that we have identified in isolation the effect of pure uncertainty on the optimal shape of contracts, we turn to the issue of risk-sharing. Indeed, one of the arguments put forward concerning relative performance evaluation is its risk-filtering property. This property is better understood when comparing relative performance evaluation to independent contracts or individual piece-rates (e.g. Lazear and Rosen, 1981) in a context when agents are risk-averse. When a common noise influences the performance of the two agents, the principal can use the output of one agents to at least partially correct for the common noise in the other agent's incentive scheme, which reduces the risk-premium to be conceded. The trade-off between incentives and insurance is then solved at smaller costs.

To treat those aspects, we consider a variant of the model which has the following features:

- Agents are risk averse with utility separable between money and effort, the monetary part is evaluated according to the concave function u.
- Agents are subject to a participation constraint with reservation utility v, instead of limited liability.

Thus the payoff of agent 1 now writes:

$$U_1(\boldsymbol{w}|e, f) = \mathbb{E}_{\boldsymbol{R}}\left[u(w_{\boldsymbol{R}})|e, f\right] - c(e) \tag{4}$$

while the incentive constraints remains formally the same as (1) and (2).

$$U_1(\boldsymbol{w}|1, 1) \geq U_1(\boldsymbol{w}|0, 1) \tag{5}$$

In turn, the limited liability constraint for agent 1 is replaced by the following participation constraint:

$$U_1(\boldsymbol{w}|1,1) \geq v \tag{6}$$

We are in position to solve the principal's problem, and to obtain a picture parallelling the results of the preceding sections. As a first step, the next lemma is a smooth equivalent to lemma 2:

Lemma 4 *Under risk-aversion, the optimal wages are ranked according to their incentive efficiency.*

Now, the optimal incentive scheme can be fully characterized:

Proposition 3 *When assumption 1 holds and the agents are risk-averse, the optimal wage profile can generically be of four different types:*

- *Relative performance evaluation*
- *Joint performance evaluation*
- *Relative bonuses: Profit sharing at the bottom, relative evaluation at the top:*

$$w_{SF} > w_{SS} > w_{FS} > w_{FF}$$

- *Relative penalties: Profit sharing at the top and relative evaluation at the bottom:*

$$w_{SS} > w_{SF} > w_{FF} > w_{FS}$$

Note that IPE can only occur at the origin in terms of covariances. That is, even if there is no correlation in equilibrium, it may be strictly optimal to use dependent compensation scheme only because out-of-equilibrium correlation exist. The interpretation of this proposition follows directly the lines of the discussion after proposition 2. In addition, the case of a failure of the agent under consideration (i.e. the cases FF and FS) matters here, since no wages are constrained by limited liability, and all four wages have to be chosen. Similarly to the case of a success, the relative informativeness ($\frac{\sigma_e}{1-p_e}$) of a failure is crucial. It is particularly interesting that *mixed* schemes are often optimal. Those schemes mixing an element of RPE and an element of JPE have clear economic interpretations and are indeed used in practice. For example, they correspond to the combination of profit sharing with selective promotions (in the first case) or selective firing (in the third case).

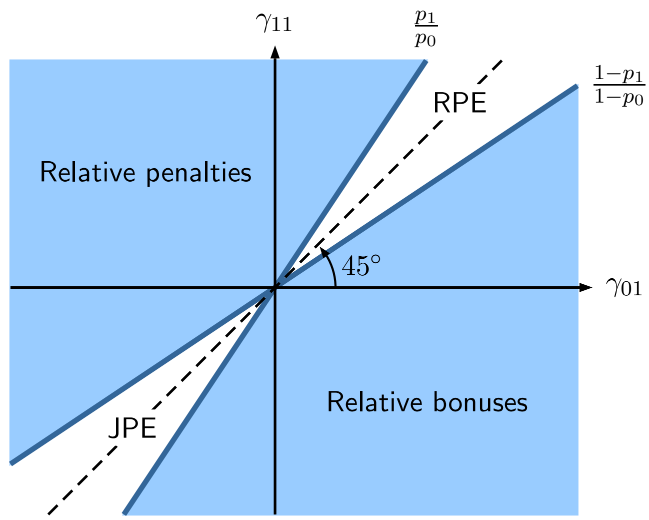

Figure 2: Optimal scheme in the covariance space with risk-aversion

Finally, this suggests that an analysis in which the level of relative evaluation is constant over results pairs[17] such as in the LEN model (Holmström and Milgrom, 1990; Itoh, 1992), see Appendix A, is unsatisfactory in that it does not allow for such mixed schemes. The full picture emerging from the proposition is summarized in the next figure.

5 Conclusion

The main message of this paper is twofold. First, relative performance evaluation is not necessarily the best informational tool. The nature and shape of uncertainty may matter even with risk-neutral agents, and a principal generically benefits from uncertainty in that case. Regarding the desirability of relative evaluation, the model demonstrates that standard results in multiagent moral hazard problems are not robust, and identifies the specificity of previous analyses. In particular, under risk-neutrality, high equilibrium correlation of the agents' performances pleads for joint performance evaluation. Second, under risk-aversion, correlated risks call for noise filtering to share risk optimally, and two opposite effects have thus to be traded off in the wage schedule. Optimal mixed schemes balancing those two effects typically exhibit real-life features that still require theoretical foundations. Interestingly, (relative) sticks and carrots are not equivalent in those

[17] While Holmström and Milgrom (1987) obtain conditions under which an optimal incentive scheme is linear in aggregate profit, there is no result stating that in a model with multiple observables–possibly from different agents–should be linear in those performances.

schemes. Another line of research worth pursuing is the properties of mixed schemes in dealing with the multiple equilibria problem inherent in joint performance evaluation and the collusion problem inherent in relative performance evaluation. Finally applications of the ideas developed here to the remuneration of top executives and fund managers are left for future research.

References

Aggarwal, R., Samwick, A., 1999. Executive Compensation, Strategic Competition, and Relative Performance Evaluation: Theory and Evidence, Journal of Finance 54(6), 1999-2043.

Anscombe, F., Aumann, R., 1963. A Definition of Subjective Probability, Annals of Mathematical Statistics 34, 199205.

d'Aspremont, C., Gérard-Varet, L.-A., 1998. Linear Inequality Methods to Enforce Partnerships under Uncertainty: An Overview, Games and Economic Behavior 25, 3SS-336.

Arya, A., Glover, J., 1995. A Simple Forecasting Mechanism for Moral Hazard Settings, Journal of Economic Theory 66, 507-521.

Arya, A., Glover, J., Hughes, J., 1997. Implementing Coordinated Team Play, Journal of Economic Theory 74, 218-232.

Baliga, S., Sjöström, T., 1998. Decentralization and Collusion, Journal of Economic Theory 83, 196-232.

Bolton, P., Harris, C., 1999. Strategic Experimentation, Econometrica 67(2), 349-374.

Brusco, S., 1997. Implementing Action Profiles when Agents Collude, Journal of Economic Theory 73, 395-424.

Che, Y.K., Yoo, S.-W., 2001. Optimal Incentives for Teams, American Economic Review 91, 525-541.

DeGroot, M., 1970. Optimal Statistical Decisions, McGraw-Hill.

Demski, J., Sappington, D., 1984. Optimal incentive contracts with multiple agents, Journal of Economic Theory 33(1), 152-171.

Dewatripont, M., Jewitt, I., Tirole, J., 1999. The Economics of Career Concerns, Part 1 and Part 2, Review of Economic Studies 66, 183-217.

Ghirardato, P., Maccheroni, F., Marinacci, M., 2004. Differentiating ambiguity and ambiguity attitude, Journal of Economic Theory 118(2), 133-173.

Gilboa, I., Schmeidler, D., 1989. Maxmin expected utility with non-unique prior, Journal of Mathematical Economics 18(2), 141-153.

Green, J., Stokey, N., 1983. A Comparison of Tournaments and Contracts, Journal of Political Economy 91(3), 349-364.

Halevy, Y., Feltkamp, V., 2005. A Bayesian Approach to Uncertainty Aversion, Review of Economic Studies 72(2), 449-466.

Holmström, B., 1979. Moral Hazard and Observability, Bell Journal of Economics 10(1), 74-91.

Holmström, B., 1982. Moral Hazard in Teams, Bell Journal of Economics 13, 324-340.

Holmström, B., Milgrom, P., 1987. Aggregation and Linearity in the Provision of Intertemporal Incentives, Econometrica, 55(2), 303-328.

Holmström, B., Milgrom, P., 1990. Regulating Trade Among Agents, Journal of Institutional and Theoretical Economics 146, 85-105.

Itoh, H., 1991. Incentives to Help in Multi-Agents Situations, Econometrica 59, 611-636.

Itoh, H., 1992. Cooperation in Hierarchical Organizations: an Incentive Perspective, Journal of Law, Economics and Organization 8, 321-345.

Itoh, H. 1993. Coalitions, Incentives and Risk Sharing, Journal of Economic Theory 60, 410-427.

Jewitt, I., 1988. Justifying the First-Order Approach to Principal-Agent Problems, Econometrica 56, 1177-1190.

Klibanoff, P., Marinacci, M., Mukerji, S., 2005. A Smooth Model of Decision Making under Ambiguity, Econometrica 73, 1849-1892.

Laffont, J.-J., Martimort, D., 2002. The Theory of Incentives, Princeton University Press, Princeton, NJ.

Lambert, R., 2001. Contracting Theory and Accounting, Journal of Accounting and Economics 32, 387.

Lazear, E., 1989. Pay Equality and Industrial Politics, Journal of Political Economy 97, 561-580.

Lazear, E., Rosen, S., 1981. Rank-Order Tournaments as Optimum Labor Contracts, Journal of Political Economy 89(5), 841-864.

Luporini, A., 2006. Relative performance evaluation in a multi-plant firm, Economic Theory 28, 235-243.

Ma, C.-T., 1988. Unique implementation of incentive contracts with many agents, Review of Economic Studies 55, 555-572.

Ma, C.-T., Moore, J., Turnbull, S., 1988. Stopping agents from "cheating", Journal of Economic Theory, 46(2), 355,372.

Maskin, E., Qian, Y., Xu, C., 2000. Incentives, Information and Organizational Form, Review of Economic Studies 67, 358-378.

Mookherjee, D., 1984. Optimal Incentives Schemes with Many Agents, Review of Economic Studies 51, 433-446.

Nalebuff, G., Stiglitz, J., 1983. Prices and Incentives: Towards a General Theory of Compensation and Competition, Bell Journal of Economics 14, 21-43.

Ramakrishnan, R., Takor, A., 1991. Cooperation versus Competition in Agency, Journal of Law, Economics and Organization 7, 248-283.

A Information Structure of the LEN model

A.1 A quick reminder

The Linear-Exponential-Normal model has been popularized by the seminal paper of Holmström and Milgrom (1987), and applied to the multiagent setting by Holmström and Milgrom (1990) and Itoh (1992) among others. It is especially popular in the accounting and management literature, see Lambert (2001, pp. 29-47) for an overview of its extensive use. This popularity can be explained by its tractability: indeed, with linear incentives schemes (the 'L' of LEN), exponential utility ('E') and normal noise ('N'), the derivation of optimal incentive coefficients amounts to a simple quadratic optimization problem.

To parallel the model of this paper, we adopt the notations e and f to denote efforts of agent 1 and 2 respectively. As is customary, we assume that the signals received by the principal are $R_1 = e + \varepsilon$ and $R_2 = f + \eta$ where (ε, η) is a joint normal distribution with variance covariance matrix $\begin{vmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{vmatrix}$. Note that symmetry is assumed for expositional clarity, but is not required.

The assumptions of linear wages imply the following incentive schemes:

$$w_1 = \alpha_1 + \beta_{11} R_1 + \beta_{12} R_2$$

$$w_2 = \alpha_2 + \beta_{21} R_1 + \beta_{22} R_2$$

In such a context, a scheme is RPE when $\beta_{ij} < 0$ for $i \neq j$, and collective if those coefficient are positive. Moreover, the magnitude of those coefficients gives a measure of the intensity of competition. The principal's optimization yields the following relationships between the optimal weights (see for example Holmström and Milgrom, 1990):

$$\beta_{ii}^* = \frac{1}{1 + r\sigma^2(1 - \rho^2)}$$

$$\beta_{ij}^* = -\rho\beta_{ii} \text{ for } i \neq j$$

Therefore, a positive correlation of the performances calls for competitive (RPE) schemes, since $\beta_{ij}^* < 0$ for $\rho > 0$. Moreover, one has $\beta_{ij}^*/\beta_{ii}^* = -\rho$, and therefore the optimal scheme is relatively more competitive when ρ is higher, a central idea in the literature.

A.2 Correlation and quality of information

A first fact regarding the LEN model is that the correlation of the performances does not varies with the effort pair (e, f). Indeed, one has:

$$corr(e + \varepsilon, f + \eta) = \rho$$

while we have seen that in the model developed above the ratio ρ_{11}/ρ_{01}, measuring how correlation varies depending on whether the two agents choose the same action or not is crucial. It seems for example unrealistic that the performance of an agent that works very hard is as correlated to the performance of an agent that does not work at all as it is correlated to the performance of an agent that works equally hard. The model developed above allows a complete flexibility with regard to that aspect.

Second, the quality of information on effort contained in the signal observed by the principal–that is, the signal to noise ratio- in the LEN model satisfies:

$$\frac{\partial}{\partial e}\left(\frac{e}{var(e + \varepsilon)}\right) > 0$$

which expresses the fact that the signal is all the more precise than effort is higher. Clearly, the magnitude of the normal noise being constant, or put differently, the noises being homoscedastic, the relative error becomes smaller when the effort becomes higher. Therefore, the LEN model does not capture situations in which more effort induces a relatively more risky situation (see in particular Dewatripont et al., 1999, for a discussion of that aspect in the career concerns model).

This overall suggests that the conclusions drawn from the LEN model are specific in that they do not give as wide a description as the present model in terms of the various potential effects of effort levels (in and out of equilibrium) on the quality of information and related optimal incentives schemes. In particular, the LEN model unreasonably favors competitive schemes, on both dimensions emphasized in the criterion of proposition 2.

B Omitted Proofs

B.1 Proof of lemma 2

In the principal's program, let $\lambda > 0$ be the Lagrange multiplier associated with the incentive constraint, and $\mu_{\mathbf{R}} \geq 0$ that associated with the limited liability constraint $w_{\mathbf{R}} \geq 0$. The first-order conditions for each $w_{\mathbf{R}}$ is:

$$-Prob(\mathbf{R}|1,1) + \lambda(Prob(\mathbf{R}|1,1) - Prob(\mathbf{R}|0,1)) + \mu_{\mathbf{R}} = 0$$

If a wage $w_{\mathbf{R}}$ is positive then $\mu_{\mathbf{R}} = 0$ and the last equation writes:

$$I(\mathbf{R}) = \frac{1}{\lambda}$$

For a wage equal to zero, say $w_{\mathbf{R'}}$, one has $\mu_{\mathbf{R'}} I(\mathbf{R'}) = \frac{1}{\lambda}(1 - \frac{\mu_{\mathbf{R'}}}{Prob(\mathbf{R'}|1,1)}) < \frac{1}{\lambda}$, hence the conclusion.

B.2 Proof of proposition 1

Consider the case of complementary effort. By definition, we have:

$$\frac{Prob(SS|1,1)}{Prob(SF|1,1)} \geq \frac{Prob(SS|0,1)}{Prob(SF|0,1)} \Leftrightarrow \frac{Prob(SS|1,1)}{Prob(SS|0,1)} \geq \frac{Prob(SF|1,1))}{Prob(SF|0,1)}$$
$$\Leftrightarrow h(SS) \geq h(SF)$$

and

$$\frac{Prob(FS|1,1)}{Prob(FF|1,1)} \geq \frac{Prob(FS|0,1)}{Prob(FF|0,1)} \Leftrightarrow \frac{Prob(FS|1,1)}{Prob(FS|0,1)} \geq \frac{Prob(FF|1,1))}{Prob(FF|0,1)}$$
$$\Leftrightarrow h(FS) \geq h(FF)$$

From lemma 2, the only wages that can be positive are thus w_{SS} and w_{FS}. Note that we used equivalences, hence the conclusion. The case of substitute is dealt with similarly.

B.3 Proof of lemma 3

By complementary probabilities and independent productions, we have the identities:

$$Prob(SS|1,1) - Prob(SS|0,1) = -\left(Prob(FS|1,1) - Prob(FS|0,1)\right)$$
$$Prob(SF|1,1) - Prob(SF|0,1) = -\left(Prob(FF|1,1) - Prob(FF|0,1)\right)$$

so that the incentive constraint can be written as:

$$(Prob(SS|1,1) - Prob(SS|0,1))(w_{SS} - w_{FS}) \\ + (Prob(SF|1,1) - Prob(SF|0,1))(w_{SF} - w_{FF}) \geq c$$

Now, we have:

$$Prob(SS|1,1) - Prob(SS|0,1) = \mathbb{E}[\tilde{p}_1\tilde{q}_1] - \mathbb{E}[\tilde{p}_0\tilde{q}_1] = \mathbb{E}[\tilde{q}_1(\tilde{p}_1 - \tilde{p}_0)]$$

From assumption 1, $(\tilde{p}_1 - \tilde{p}_0)$ is a positive random variable, as is $\tilde{q}_1$. Thus the coefficient of w_{FS} in the incentive constraint is negative, which implies that this wage should be 0. Similarly, one has:

$$\begin{aligned} Prob(SF|1,1) - Prob(SF|0,1) &= \mathbb{E}[\tilde{p}_1(1-\tilde{q}_1)] - \mathbb{E}[\tilde{p}_0(1-\tilde{q}_1)] \\ &= \mathbb{E}[(1-\tilde{q}_1)(\tilde{p}_1 - \tilde{p}_0)] \end{aligned}$$

which is also positive from the assumption.

B.4 Proof of proposition 2

From the two preceding lemmata, we know that except in the special case $I(SS) = I(SF)$ only one wage is positive. The criterion for $w_{SS} > 0$ is $I(SS) > I(SF)$. We need the following simple calculation to undertake the comparison:

$$\begin{aligned} Prob(SS|11) &= \mathbb{E}[\tilde{p}_1\tilde{q}_1] = p_1q_1 + \rho_{11}\sigma_1\tau_1 \\ Prob(SS|01) &= \mathbb{E}[\tilde{p}_0\tilde{q}_1] = p_0q_1 + \rho_{01}\sigma_0\tau_1 \\ Prob(SF|11) &= \mathbb{E}[\tilde{p}_1(1-\tilde{q}_1)] = p_1(1-q_1) - \rho_{11}\sigma_1\tau_1 \\ Prob(SF|01) &= \mathbb{E}[\tilde{p}_0(1-\tilde{q}_1)] = p_0(1-q_1) - \rho_{01}\sigma_0\tau_1 \end{aligned}$$

Using those values yields:

$$\frac{Prob(SS|1,1)}{Prob(SS|0,1)} > \frac{Prob(SF|1,1)}{Prob(SF|0,1)} \Leftrightarrow \frac{q_1p_1 + \rho_{11}\sigma_1\tau_1}{p_0q_1 + \rho_{01}\sigma_0\tau_1} > \frac{p_1(1-q_1) - \rho_{11}\sigma_1\tau_1}{p_0(1-q_1) - \rho_{01}\sigma_0\tau_1}$$

which simply boils down to

$$\rho_{11} > \rho_{01}\frac{p_1}{\sigma_1}\frac{\sigma_0}{p_0}$$

Conversely, one easily obtains that w_{SF} is positive under the reverse inequality. The optimal wages are then straightforwardly obtained by saturating the incentive constraint. In the case of equality, both wages have the same incentive weight, and only their sum matters. The optimal sum is also obtained by saturating the incentive constraint.

B.5 Proof of lemma 4

We associate the positive multipliers $\lambda \geq 0$ and $\mu \geq 0$ to, respectively, the incentive and participation constraints, and form the Lagrangian of the cost minimization problem:

$$\begin{aligned} L(\boldsymbol{w}, \lambda, \mu) &= \Sigma_{\boldsymbol{R}} Prob(\boldsymbol{R}|11) w_{\boldsymbol{R}} + \lambda c + \mu(v + c) \\ &\quad - \Sigma_{\boldsymbol{R}} \left[\lambda \left(Prob(\boldsymbol{R}|11) - Prob(\boldsymbol{R}|01)\right) + \mu Prob(\boldsymbol{R}|11)\right] u(w_{\boldsymbol{R}}) \end{aligned}$$

It is clear that both multipliers have to be positive. The first-order conditions for all $\boldsymbol{R}$ boil down to:

$$\frac{1}{u'(w_{\boldsymbol{R}})} = \mu + \lambda \frac{Prob(\boldsymbol{R}|11) - Prob(\boldsymbol{R}|01)}{Prob(\boldsymbol{R}|11)} = \mu + \lambda I(\boldsymbol{R})$$

Note that $\frac{1}{u'}$ is an increasing function, thus w's are ranked as $\frac{1}{u'(w)}$. This means that the wages are ranked according to their incentive efficiency.

B.6 Proof of proposition 3

The ranking between w_{SS} and w_{SF} corresponds to the criterion of the first proposition. Also, from assumption 1, we have:

$$\begin{aligned} Prob(FS|11) - Prob(FS|01) &= \mathbb{E}[(1 - \tilde{p}_1)\tilde{q}_1] - \mathbb{E}[(1 - \tilde{p}_0)\tilde{q}_1] \\ &= \mathbb{E}[\tilde{q}_1(\tilde{p}_0 - \tilde{p}_1)] \leq 0 \end{aligned}$$

and

$$\begin{aligned} Prob(FF|11) - Prob(FF|01) &= \mathbb{E}[(1 - \tilde{p}_1)(1 - \tilde{q}_1)] - \mathbb{E}[(1 - \tilde{p}_0)(1 - \tilde{q}_1)] \\ &= \mathbb{E}[(1 - \tilde{q}_1)(\tilde{p}_0 - \tilde{p}_1)] \leq 0 \end{aligned}$$

Which indicate that both $I(FS)$ and $I(FF)$ are negative, while we have already seen that $I(SS)$ and $I(SF)$ are positive. This implies that w_{SS} and w_{SF} are always higher than w_{FF} and w_{FS}. To finish the proof, we need the ranking between w_{FF} and w_{FS}, which requires a few additional calculations:

$$\begin{aligned} Prob(FS|11) &= \mathbb{E}[(1 - \tilde{p}_1)\tilde{q}_1] = q_1(1 - p_1) - \rho_{11}\sigma_1\tau_1 \\ Prob(FS|01) &= \mathbb{E}[(1 - \tilde{p}_0)\tilde{q}_1] = q_1(1 - p_0) - \rho_{01}\sigma_0\tau_1 \\ Prob(FF|11) &= \mathbb{E}[(1 - \tilde{p}_1)(1 - \tilde{q}_1)] = (1 - p_1)(1 - q_1) + \rho_{11}\sigma_1\tau_1 \\ Prob(FF|01) &= \mathbb{E}[(1 - \tilde{p}_0)(1 - \tilde{q}_1)] = (1 - p_0)(1 - q_1) + \rho_{01}\sigma_0\tau_1 \end{aligned}$$

Using those values and conducing calculations parallelling that in the other comparison yields:

$$I(FS) - I(FF) > 0 \Leftrightarrow \rho_{11} < \rho_{01} \frac{\sigma_0}{1 - p_0} \frac{1 - p_1}{\sigma_1}$$

Note that a JPE scheme can now never be optimal for positive correlation, since it would require at the same time $w_{SS} \geq w_{SF}$ and $w_{FS} \geq w_{FF}$. All the other combinations are in turn possible, depending on the parameters. When both covariances are negative, RPE is excluded, and the analysis is virtually the same.