Annals of Finance (2006)
DOI 10.1007/s10436-006-0044-0

RESEARCH ARTICLE

Michael Magill · Martine Quinzii

Common shocks and relative compensation

Received: 6 January 2006 / Accepted: 4 March 2006

Abstract This paper studies qualitative properties of an optimal contract in a multi-agent setting in which agents are subject to a common shock. We derive a necessary and sufficient condition for the optimal reward of an agent producing an output level y to be a decreasing (increasing) function of the outputs of the other agents, under the assumption that the agents' outputs are informative signals of the value of the common shock. The condition is that the likelihood ratio $p(y, e, \eta)/p(y, e', \eta)$, where e is a higher effort level than e' and η is the value of the common shock, be a decreasing (increasing) function of η. We give examples of applications of the result and examine its consequences for CEO compensation.

Keywords Optimal contracts · Reward increasing (decreasing) in other agents' outcomes · Likelihood ratio and common shock · Effect of common shock on marginal product of effort

JEL Classification Numbers D82 · G30 · J33 · M52

1 Introduction

It has been shown by Holmström (1982) and Mookherjee (1984) that when a common shock affects the performance of several agents, the optimal contract of one agent depends on the performance of the others. Holmström also showed that under specific assumptions on the production function and under normality assumptions

M. Magill
Department of Economics, University of Southern California, Los Angeles,
CA 90089-0253, USA
E-mail: magill@usc.edu

M. Quinzii (✉)
Department of Economics, University of California, Davis,
CA 95616-8578, USA
E-mail: mmquinzii@ucdavis.edu

on the distribution of the shocks, the information provided by the performance of the other agents can be summarized in an average, which is a sufficient statistic for the common shock. The usual interpretation of this result, based on specific examples of production functions involving effort, idiosyncratic shocks, and an additive common shock, is that the optimal contract of an agent should use the performance of other agents to "factor out" the effect of the common shock on the observed performance of the agent. According to this theory of *relative performance evaluation*, for a given outcome of an agent, the compensation should be lower if other agents have on average a good outcome, since this indicates that the shock has been favorable: a good outcome for the agent is then less likely to be attributable to high effort, while a bad outcome makes low effort on the part of the agent more likely.

The design of compensation packages of CEOs of corporations is a natural application of the theory of optimal multi-agent contracts when outcomes are observable and are affected by a common shock. For there is a high observed correlation among firms' outcomes, and this is usually attributed to the fact that firms are subject to common sectoral and/or economy wide shocks. However, if the observed compensation packages of CEOs include some degree of relative performance evaluation in the bonus components of their pay, in the 1990s options became the largest component of CEO compensation, and no attempt was made to index their exercise prices to "factor out" the general level of the market (Murphy 1999; Himmelberg and Hubbard 2000). As a result, during the bull market of the 1990s, the compensation of CEOs increased with the average performance of the firms in the economy, rather than with their relative performance. This is viewed by many authors in corporate finance as an indication that the compensation packages designed by boards of directors are inefficient.

In the standard literature the intuition that an optimal contract should factor out the effect of a common shock is based on examples with specific production functions in which the common shock enters additively (Lazear and Rosen 1981; Green and Stokey 1983). In fact little has been established in a general setting on the way an optimal contract should use the information provided by the realized performance of other agents. Should the reward of an agent always decrease when the performance of other agents in the comparison group increases, or can there be circumstances when it is optimal that it increase?

In this paper we derive a necessary and sufficient condition for the optimal reward of an agent producing an output level y to be a decreasing (increasing) function of the outputs of the other agents, under the assumption that the agents' outputs are informative signals of the value of the common shock. The condition (expressed in the paper in its differential form) is that the likelihood ratio $p(y, e, \eta)/p(y, e', \eta)$, where e is a higher effort level than e', and η is the value of the common shock, be a decreasing (increasing) function of η. If y is a high outcome, a decreasing likelihood ratio formalizes the idea that the more favorable the common shock, the less likely it is that the observed output y is attributable to high rather than low effort, while if y is a low outcome, the more likely it is that y is to be attributed to low effort. According to the principle that an incentive contract should reward an agent in circumstances which are likely to occur when effort is high, and punish the agent in circumstances which are likely when effort is low, the compensation decreases when the performance of other agents increases.

When the likelihood ratio is increasing rather than decreasing in η, the reward of an agent increases with the performance of the other agents.

In Sect. 4 we derive conditions under which the likelihood ratio is decreasing for all output levels, or increasing for some output levels and decreasing for others. The conditions hinge on the way the common shock affects the marginal product of effort. If the shock enters additively and does not affect the marginal product of effort, as in the model of Green and Stokey (1983), then the optimal contract is 'tournament-like' in that the payoff of an agent always decreases when the performance of other agents increases. When the common shock positively affects the productivity of effort, as in the model of Nalebuff and Stiglitz (1983), a higher shock tends to raise the productivity of effort. In this case sufficiently high outcomes are more likely to come from high effort, while low outcomes are always more likely to come from low effort: thus for low outcomes the reward is a decreasing function of the performance of others, while for sufficiently high outcomes it is increasing. When the shock adversely affects the productivity of effort, the effects are reversed.

Recently several authors, motivated by observed CEO compensations in the 1990s, have proposed models in which the optimal contract is such that, in some circumstances, the CEO is paid more if the common shock is favorable. One such model is presented by Celentani and Loveira-Pazo (2004). Although their model is based on adverse selection rather than moral hazard, the condition that they find is close in spirit to the increasing likelihood ratio property: they show that the optimal contract has the desired feature that the payment in case of a good outcome is higher when the outside shock is favorable if "success is relatively more likely to derive from the good manager rather than the bad one" in the good state than in the bad state. The model which is closest to our framework is the model proposed by Himmelberg and Hubbard (2000). They specify the production and utility function of a standard principal agent model in order to apply it to CEO compensation: in their specification the marginal product of effort of the CEO increases with the value of the aggregate shock, a feature that they suggest is realistic for CEOs.

The model is outlined in Sect. 2 and the main result is established in Sect. 3. Examples illustrating the result are given in Sect. 4, while the final Section concludes with some remarks on executive compensation suggested by our analysis.

2 Model and assumptions

Consider a collection of K firms which produce a homogenous output (profit) and a collection of K managers who run these firms. Managers are assumed to be matched to firms: manager[1] k in $\mathcal{K}$ can only manage firm k or take an outside option which determines his reservation utility $\bar{v}_k$ for working for firm k. The output y^k of firm k depends on the entrepreneurial effort e_k of its manager, on a random shock $\eta \in \Re$ which is common to all firms and on an idiosyncratic shock ϵ_k, where both types of shocks are unobserservable: thus $y^k = h^k(e_k, \eta, \epsilon_k)$. We assume that effort $e_k \in \Re_+$ is a continuous variable and that the output levels h^k take a finite number

[1] We use a caligraphic letter to denote a set and the same roman uppercase letter to denote the number of its elements: $\mathcal{K} = \{1, \dots, K\}$

of values[2] indexed in increasing order by $s_k \in \mathcal{S}_k$: that is $s_k > s'_k \Longrightarrow y^k_{s_k} > y^k_{s'_k}$. For given e_k and η, the distribution function of ϵ_k induces a probability distribution $p_k(\cdot, e_k, \eta)$ on $\mathcal{S}_k$ whose cumulative distribution function is denoted by F_k: that is, $F_k(\alpha, e_k, \eta) \equiv \sum_{\{s_k | y^k_{s_k} \leq \alpha\}} p_k(s_k, e_k, \eta)$. We assume that $p_k(s_k, e_k, \eta)$ is a differentiable function of e_k.

Let $\mathcal{S} = \mathcal{S}_1 \times \cdots \times \mathcal{S}_K$. An outcome for the economy is a realization $\boldsymbol{s} = (s_1, \ldots, s_K) \in \mathcal{S}$, namely a vector of realized outputs $\boldsymbol{y_s} = (y^1_{s_1}, \ldots, y^K_{s_K})$ for the K firms. When we consider the optimal contract for manager k it will be convenient to use the notation $\boldsymbol{s} = (s_k, \boldsymbol{s}^{-k})$, where $\boldsymbol{s}^{-k} = (s_1, \ldots, s_{k-1}, s_{k+1}, \ldots, s_K)$, and similarly $\boldsymbol{e} = (e_k, \boldsymbol{e}^{-k})$ for the vector of effort levels of the managers. For a given vector $\boldsymbol{e} = (e_1, \ldots, e_K)$ and a given value of the common shock η, the probability of observing outcome $\boldsymbol{s}$ is

$$p(\boldsymbol{s}, \boldsymbol{e}, \eta) \equiv \prod_{k \in \mathcal{K}} p_k(s_k, e_k, \eta),$$

i.e. the idiosyncratic shocks $(\epsilon_1, \ldots, \epsilon_K)$ are assumed to be independent, conditional on the value of η. Let $G(\eta)$ denote the distribution function of η; since η is unobservable the probability of outcome $\boldsymbol{s}$ given $\boldsymbol{e}$ is

$$P(\boldsymbol{s}, \boldsymbol{e}) \equiv \int_{\Re} p(\boldsymbol{s}, \boldsymbol{e}, \eta) \mathrm{d}G(\eta).$$

The random variables $\{\epsilon_1, \cdots, \epsilon_K, \eta\}$ are not observable by any agent but the uncertainty structure $(p_1, \ldots, p_K, G)$ of the economy is assumed to be common knowledge.

Each firm k is owned by a collection of risk-neutral shareholders, who hire the manager and offer an incentive contract $\boldsymbol{\tau}_k = (\tau_k(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}) \in \Re^S$, which guarantees the manager the reservation utility level $\bar{v}_k$. Each manager is assumed to be risk averse and, given the pay schedule $\boldsymbol{\tau}_k$ and effort e_k, has utility

$$U_k(\boldsymbol{\tau}_k, e_k) = E(u_k(\boldsymbol{\tau}_k)) - c_k(e_k),$$

where $u_k, c_k : \Re_+ \to \Re$ are differentiable, increasing, u_k is strictly concave and c_k is convex.

Because the realization $\boldsymbol{s}^{-k}$ of firms other than k contains information about the common shock η, the optimal contract $\boldsymbol{\tau}_k$ for manager k will use this information to provide incentives at least cost for the shareholders. The contract $\boldsymbol{\tau}_k$ will thus depend on the realized state $s = (s_k, \boldsymbol{s}^{-k})$ and since the probability of the realization $\boldsymbol{s}^{-k}$ depends on the effort levels $\boldsymbol{e}^{-k}$ of the other agents, the optimal effort of manager k will indirectly depend on the effort levels of the other managers.

[2] The principal agent model is studied either with discrete effort levels as in Grossman and Hart (1983) or with a continuum of effort levels as in most of the papers cited in the introduction. Here we study the case of a continuum of effort levels and use first-order conditions with respect to effort in the proof of Proposition 1. A modified proof is needed to cover the case of discrete effort levels. As for the outcomes of the firms, we take the case of discrete outcomes since the notation and presentation of the model is simpler and more intuitive in this case. It is easy to see that the proof of Proposition 1 carries over to the case of continuous outcomes, integrals replacing sums in the obvious way.

The assumption that the firms' owners are risk neutral, while standard in the principal-agent literature, is nevertheless restrictive. The usual justification is that if investors are well diversified and own a sufficiently small share of each firm, their risk aversion can be neglected in designing the optimal contract of a firm's manager. In this paper the assumption of risk neutrality is necessary to obtain clear-cut results on the way the optimal contract of a firm's manager depends on the outcomes of other firms. For if investors were assumed to be risk averse, the contract of manager k would depend on the outcomes of other firms for two reasons: first because all agents would have to share the aggregate risks, and second because the outcomes of other firms give information on the likely effort of manager k. By assuming risk neutrality of the investors we eliminate the need to consider the first effect—namely the need to share risk between investors and managers—so that the second effect—namely the informational content of the outcomes of other firms—is the only reason for which they enter into the manager's contract.

We will restrict attention to interior Nash equilibria in which the shareholders of each firm induce the manager to make a positive effort, and each manager receives a positive payment in every realized outcome.

Definition $(\bar{\boldsymbol{\tau}}, \bar{\boldsymbol{e}}) = ((\bar{\boldsymbol{\tau}}_1, \ldots, \bar{\boldsymbol{\tau}}_K), (\bar{e}_1, \ldots, \bar{e}_K)) \gg 0$ *is an interior Nash equilibrium with optimal incentive contracts if for each* $k \in \mathcal{K}$, $(\bar{\boldsymbol{\tau}}_k, \bar{e}_k)$ *solves the profit maximization problem*

$$\max_{(\boldsymbol{\tau}_k, e_k) \in \Re_+^S \times \Re_+} \sum_{\boldsymbol{s} \in \mathcal{S}} P\left(\boldsymbol{s}, e_k, \bar{\boldsymbol{e}}^{-k}\right) \left(y_{s_k}^k - \tau_k(\boldsymbol{s})\right)$$

subject to

$$\sum_{\boldsymbol{s} \in \mathcal{S}} P\left(\boldsymbol{s}, e_k, \bar{\boldsymbol{e}}^{-k}\right) u_k(\tau_k(\boldsymbol{s})) - c_k(e_k) \geq \bar{v}_k \tag{1}$$

$$e_k \in \arg\max \left\{ \sum_{\boldsymbol{s} \in \mathcal{S}} P(\boldsymbol{s}, \tilde{e}_k, \bar{\boldsymbol{e}}^{-k}) u_k(\tau_k(\boldsymbol{s})) - c_k(\tilde{e}_k) | \tilde{e}_k \in \Re_+ \right\}. \tag{2}$$

The main result presented in the next section is based on the analysis of the first-order conditions for the profit maximizing problem of each firm at a Nash equilibrium. Given the form of the incentive constraints (2), these FOCs cannot be derived directly from the Kuhn-Tucker theorem. We derive them using the following assumption.

(A1) $(\bar{\boldsymbol{\tau}}, \bar{\boldsymbol{e}})$ is an interior Nash equilibrium such that, for each $k \in \mathcal{K}$, $\bar{e}_k$ is the unique optimal effort of manager k given the compensation schedule $\bar{\boldsymbol{\tau}}_k$.

The uniqueness condition is generic: if $(\bar{\boldsymbol{\tau}}, \bar{\boldsymbol{e}})$ is an equilibrium such that the maximum is not unique for some manager k, i.e. $U_k(\bar{\boldsymbol{\tau}}_k, \tilde{e}_k) = U_k(\bar{\boldsymbol{\tau}}_k, \bar{e}_k)$ for $\tilde{e}_k \neq \bar{e}_k$, then the cost function c_k can be perturbed so that $\bar{e}_k$ becomes the unique optimal effort without affecting the other agents, and $(\bar{\boldsymbol{\tau}}, \bar{\boldsymbol{e}})$ is a Nash equilibrium of the perturbed economy. (A1) is a weaker condition than the concavity of the function $e_k \rightarrow U_k(\bar{\boldsymbol{\tau}}_k, e_k)$ which is required for the "first-order approach" to hold,[3] i.e. for replacing the incentive constraint (2) by the first-order condition for optimal effort for manager k.

Lemma 1 *If (A1) holds then for each $k \in \mathcal{K}$ there exists multipliers $(\lambda_k, \mu_k) \in \Re^2_{++}$ such that*

$$1 = \left(\lambda_k + \mu_k \frac{\frac{\partial}{\partial e_k} P(\boldsymbol{s}, \bar{\boldsymbol{e}})}{P(\boldsymbol{s}, \bar{\boldsymbol{e}})} \right) u_k'(\bar{\tau}^k(\boldsymbol{s})), \quad \boldsymbol{s} \in \mathcal{S} \tag{3}$$

Proof see Appendix.

Our second assumption specifies the qualitative way in which the common shock η affects the probability of firms' outcomes.

(A2) For each $k \in \mathcal{K}$ and for any fixed $e_k \in \Re_+$, $p_k(\cdot, e_k, \cdot) : \mathcal{S}_k \times \Re \rightarrow [0, 1]$ is log-supermodular,[4] i.e., for any $\eta, \eta' \in \Re$ with $\eta > \eta'$, the likelihood ratio $p_k(s_k, e_k, \eta)/p_k(s_k, e_k, \eta')$ is an increasing function of s_k.

(A2) implies that the higher the realization s_k of firm k, the more likely it is that it is associated with the value η of the common shock rather than with the lower value η'. Thus a high realization s_k can be interpreted as a signal that the value of the common shock has been high (Milgrom 1981). Log-supermodularity is a symmetric property: if $s_k > s_k'$ and $\eta > \eta'$, then

$$\frac{p_k(s_k, e_k, \eta)}{p_k(s_k, e_k, \eta')} > \frac{p_k(s_k', e_k, \eta)}{p_k(s_k', e_k, \eta')} \iff \frac{p_k(s_k, e_k, \eta)}{p_k(s_k', e_k, \eta)} > \frac{p_k(s_k, e_k, \eta')}{p_k(s_k', e_k, \eta')}$$

so that for $s_k > s_k'$ the likelihood ratio is an increasing function of η: a higher realization of η makes it more likely that s_k rather than s_k' will occur. Since (A2) implies that a higher value of η makes larger outcomes for each firm more likely, the distribution function $F_k(\alpha, e_k, \eta)$ first-order stochastically dominates $F_k(\alpha, e_k, \eta')$ (see Rogerson 1985). High values of the common shock create good times for the economy, while a low value of η is a negative shock for the economy.

[3] In the setting of a single-principal, single-agent problem, Sinclair-Desgagné (1994) gives sufficient conditions on the function $P(\boldsymbol{s}, e)$, where $\boldsymbol{s}$ is a multi-dimensional signal and e is the (scalar) effort of the agent, which imply that the agent's utility function $U(\bar{\boldsymbol{\tau}}, e)$ is concave in e at the optimal contract $\bar{\boldsymbol{\tau}}$. These conditions, however, are not directly applicable in our setting since one of them requires that a high value of any component of the multi-dimensional signal is indicative of a high effort of the agent, i.e. the joint probability $P(\boldsymbol{s}, e)$ is assumed to have the Monotone Likelihood Property. In this paper we cover not only the case where a higher value of $\boldsymbol{s}^{-k}$ increases the likelihood of a high effort of manager k, but also the case where a lower value of $\boldsymbol{s}^{-k}$ increases this likelihood. Sinclair-Desgagné 's analysis could probably be extended to cover this case as well but this is outside the scope of our paper. To carry out our analysis we only need the necessity of the first-order condition for optimal effort which, under Assumption (A1), implies that the optimal contract satisfies the first-order condition (3).

[4] For $\boldsymbol{x}, \boldsymbol{y} \in \Re^n$, let $\boldsymbol{x} \vee \boldsymbol{y}$ =inf $\{\boldsymbol{z} \in \Re^n | \boldsymbol{z} \geq \boldsymbol{x}, \boldsymbol{z} \geq \boldsymbol{y}\}$ and let $\boldsymbol{x} \wedge \boldsymbol{y}$=sup$\{\boldsymbol{z} \in \Re^n | \boldsymbol{z} \leq \boldsymbol{x}, \boldsymbol{z} \leq \boldsymbol{y}\}$. A function $f : \Re^n \rightarrow \Re$ is *supermodular* if $f(\boldsymbol{x} \vee \boldsymbol{y}) + f(\boldsymbol{x} \wedge \boldsymbol{y}) \geq f(\boldsymbol{x}) + f(\boldsymbol{y})$ for all $\boldsymbol{x}, \boldsymbol{y} \in \Re^n$; f is *submodular* if the inequality is reversed. A function $h : \Re^n \rightarrow \Re_+$ is *log-supermodular* if $f = \log h$ is supermodular, i.e. $h(\boldsymbol{x} \vee \boldsymbol{y})h(\boldsymbol{x} \wedge \boldsymbol{y}) \geq h(\boldsymbol{x})h(\boldsymbol{y})$ for all $\boldsymbol{x}, \boldsymbol{y} \in \Re^n$. Here we assume that if $\boldsymbol{x} \neq \boldsymbol{y}$ the inequality is strict.

3 Result

Our main result gives conditions under which, for a given realization s_k of firm k, the compensation of manager k at equilibrium is a decreasing (or increasing) function of the vector of realizations $\boldsymbol{s}^{-k}$ of other firms. To express this property we say that a function $f : \Re^n \to \Re$ is decreasing [increasing] if for any $\boldsymbol{x}, \boldsymbol{y} \in \Re^n$ satisfying $\boldsymbol{x} \geq \boldsymbol{y},\ \boldsymbol{x} \neq \boldsymbol{y}$, we have $f(\boldsymbol{x}) < f(\boldsymbol{y})$ [$f(\boldsymbol{x}) > f(\boldsymbol{y})$]. We show that the compensation schedule of manager k is monotone in $\boldsymbol{s}^{-k}$ if the *local likelihood function* $L_k : S_k \times \Re_+ \times \Re \to \Re$ defined by

$$L_k(s_k, e_k, \eta) = \frac{\frac{\partial}{\partial e_k} p_k(s_k, e_k, \eta)}{p_k(s_k, e_k, \eta)}$$

is monotone in η.

Proposition 1 *Let (A1)–(A2) be satisfied. For each $k \in \mathcal{K}$ and any realization $s_k \in \mathcal{S}_k$, the optimal reward schedule $\bar{\tau}_k\left(s_k, \boldsymbol{s}^{-k}\right)$ in a Nash equilibrium is a decreasing (increasing) function of $\boldsymbol{s}^{-k}$ for all distribution functions $G(\eta)$ if and only if the local likelihood function $L_k(s_k, \bar{e}_k, \eta)$ is a decreasing (increasing) function of η.*

Proof ($\Leftarrow$) Suppose $L_k(s_k, \bar{e}_k, \eta)$ is decreasing in η. We want to show that if $\boldsymbol{s}, \boldsymbol{s}' \in \mathcal{S}$ are such that $s_k = s'_k$ and $s_j \geq s'_j$ for all $j \neq k$ with at least one strict inequality, then $\bar{\tau}_k(\boldsymbol{s}) < \bar{\tau}_k(\boldsymbol{s}')$. Since by Lemma 1 the first-order condition (3) holds with $\mu_k > 0$, and since u'_k is strictly decreasing, we need to show that $A < 0$, where A is defined by

$$A \equiv \frac{\partial/\partial e_k(P(\boldsymbol{s}, \bar{\boldsymbol{e}}))}{P(\boldsymbol{s}, \bar{\boldsymbol{e}})} - \frac{\partial/\partial e_k(P(\boldsymbol{s}', \bar{\boldsymbol{e}}))}{P(\boldsymbol{s}', \bar{\boldsymbol{e}})} \tag{4}$$

Note that

$$\frac{\frac{\partial}{\partial e_k} P(\boldsymbol{s}, \bar{\boldsymbol{e}})}{P(\boldsymbol{s}, \bar{\boldsymbol{e}})} = \int_{\Re} L_k(s_k, \bar{e}_k, \eta) a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta) \mathrm{d}G(\eta),$$

where

$$a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta) = \frac{\prod_{j \in \mathcal{K}} p_j(s_j, \bar{e}_j, \eta)}{\int_{\Re} \prod_{j \in \mathcal{K}} p_j(s_j, \bar{e}_j, \eta) \mathrm{d}G(\eta)}.$$

For all $\boldsymbol{s} \in \mathcal{S}$, $a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta) > 0$, $\int_{\Re} a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta) \mathrm{d}G(\eta) = 1$, and for all $\boldsymbol{s}, \boldsymbol{s}' \in \mathcal{S}$

$$\frac{a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta)}{a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \eta)} = \prod_{j \in \mathcal{K}} \frac{p_j(s_j, \bar{e}_j, \eta)}{p_j(s'_j, \bar{e}_j, \eta)} \frac{P(\boldsymbol{s}', \bar{\boldsymbol{e}})}{P(\boldsymbol{s}, \bar{\boldsymbol{e}})}.$$

By (A2), since log-supermodularity is symmetric in (s_k, η), if $s_j > s'_j$, the ratio $p_j(s_j, \bar{e}_j, \eta)/p_j(s'_j, \bar{e}_j, \eta)$ is an increasing function of η. Since $s_j > s'_j$

for at least one firm, it follows that the ratio $\lambda(\eta) \equiv a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta)/a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \eta)$ is an increasing function of η. Since $\int_{\Re} a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta)\mathrm{d}G(\eta) = \int_{\Re} \lambda(\eta) a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \eta)\mathrm{d}G(\eta) = \int_{\Re} a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \eta)\mathrm{d}G(\eta) = 1$, $\lambda(\eta)$ cannot always be strictly larger or strictly smaller than 1. Thus there exists $\bar{\eta} \in \Re$ such that $\lambda(\eta) \leq 1$ if $\eta \leq \bar{\eta}$ and $\lambda(\eta) > 1$ if $\eta > \bar{\eta}$, and $\int_{\eta \leq \bar{\eta}} \mathrm{d}G(\eta) > 0$, $\int_{\eta > \bar{\eta}} \mathrm{d}G(\eta) > 0$.

$$A = \int_{\eta \leq \bar{\eta}} L_k(s_k, \bar{e}_k, \eta) \left(a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta) - a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \eta)\right) \mathrm{d}G(\eta) + \int_{\eta > \bar{\eta}} L_k(s_k, \bar{e}_k, \eta) \left(a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta) - a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \eta)\right) \mathrm{d}G(\eta).$$

If $\eta \leq \bar{\eta}$ then $a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta) - a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \eta) \leq 0$ and since the likelihood function is a decreasing function of η, $L_k(s_k, \bar{e}_k, \eta) \geq L_k(s_k, \bar{e}_k, \bar{\eta})$ so that

$$L_k(s_k, \bar{e}_k, \eta) \left(a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta) - a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \eta)\right) \leq L_k(s_k, \bar{e}_k, \bar{\eta}) \left(a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta) - a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \eta)\right) \tag{5}$$

If $\eta > \bar{\eta}$, then $a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta) - a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \eta) > 0$ and $L_k(s_k, \bar{e}_k, \eta) < L_k(s_k, \bar{e}_k, \bar{\eta})$ so that (5) is satisfied with a strict inequality. Thus

$$A < L_k(s_k, \bar{e}_k, \bar{\eta}) \int_{\Re} \left(a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta) - a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \eta)\right) \mathrm{d}G(\eta) = 0.$$

If the function $L_k(s_k, \bar{e}_k, \cdot)$ is increasing in η then inequality (5) is reversed and $A > 0$, so that the optimal wage schedule is increasing in $\boldsymbol{s}^{-k}$.

($\Rightarrow$) Suppose $L_k(s_k, \bar{e}_k, \cdot)$ is not decreasing. Then there exist $\eta > \eta'$ such that $L_k(s_k, \bar{e}_k, \eta) \geq L_k(s_k, \bar{e}_k, \eta')$. Consider a distribution function G which puts weight only on η and η'. Since $a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \cdot)/a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \cdot)$ is increasing in η, and $\int_{\Re} a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta)\mathrm{d}G(\eta) = \int_{\Re} a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \eta)\mathrm{d}G(\eta) = 1$, it follows that $a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta') - a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \eta') < 0$ and $a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta) - a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \eta) > 0$. Thus A defined in (4) is such that

$$\begin{aligned} A \geq\ & L_k(s_k, \bar{e}_k, \eta') \left(a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta') - a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \eta'))G(\eta')\right. \\ & \left. +(a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta) - a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \eta))(1 - G(\eta'))\right) = 0 \end{aligned}$$

and, for the distribution function G, the payoff is non-decreasing in $\boldsymbol{s}^{-k}$. Thus the payoff is decreasing in $\boldsymbol{s}^{-k}$ for all distribution G only if the local likelihood function L_k is decreasing in η. □

Remark. Since $\mu_k > 0$, the first-order condition (3) implies that the higher the ratio $\frac{\partial P(\boldsymbol{s}, \bar{\boldsymbol{e}})/\partial e_k}{P(\boldsymbol{s}, \bar{\boldsymbol{e}})}$, the greater the compensation of manager k in outcome $\boldsymbol{s}$. In this expression η has been integrated out since it is not observable. If η were observable then the reward of manager k would only depend on the outcome s_k of his own firm and on η: his reward would vary with the local likelihood ratio $L_k(s_k, \bar{e}_k, \eta) = (\partial p_k(s_k, \bar{e}_k, \eta)/\partial e_k)/p(s_k, \bar{e}_k, \eta)$. Since

$$\frac{p_k(s_k, e_k, \eta)}{p_k(s_k, e_k', \eta)} = \exp \int_{e_k'}^{e_k} L_k(s_k, t, \eta)\mathrm{d}t$$

the greater the local likelihood ratio the greater the probability that s_k is associated with the higher effort e_k than with the lower effort e'_k. In the standard principal agent setting, it is usually assumed that $L_k(s_k, e_k, \eta)$ is increasing in s_k, a property called the Monotone Likelihood Ratio Condition (MLRC) which implies that higher outcomes s_k have a greater relative likelihood of being associated with the higher effort e_k rather than the lower effort e'_k. MLRC implies that the optimal reward schedule is increasing in s_k. In this paper we do not need this conditon since we study how the reward $\tau_k(s_k, \boldsymbol{s}^{-k})$ of the manager varies as a function of the outcomes $\boldsymbol{s}^{-k}$ of other firms for fixed s_k. What we need are two conditions: (A2), which permits the observation of the realizations $\boldsymbol{s}^{-k}$ of firms other than k to give an informative signal on the value of the common shock η, and the monotonicity of the local likelihood function L_k in η, which implies that the value of η gives information on the likelihood that the effort of the manager has been high rather than low. [5]

When L_k is decreasing in η, the manager's effort e_k and the common shock η are in essence "substitutes" for creating good outcomes, since a high value of η makes it less likely that an outcome s_k is associated with a high rather than a low value of effort. As a result, for a given outcome s_k of firm k, the reward of manager k decreases when the outcomes of other firms increase. When L_k is increasing in η, the manager's effort and η are "complements": a high value of η makes it more likely that an outcome s_k is associated with a high rather than a low value of effort and the reward of manager k increases when the outcomes of other firms increase. The examples in the next section illustrate natural settings where these two cases can arise.

4 Examples

We give two examples of settings where Proposition 1 can be used to analyze properties of the optimal reward schedule of a manager at an equilibrium.

Example 1 Consider the simplest symmetric setting where the characteristics of all firms and managers are the same and each firm has only two outcomes ($S_k = 2$, $k \in K$), a good outcome y_g and a bad outcome y_b, with $0 < y_b < y_g$. The optimal reward schedule for manager k is of the form $\tau_k(s_k, \boldsymbol{s}^{-k}) = \tau_k(s_k, n(\boldsymbol{s}^{-k}))$ where $n(\boldsymbol{s}^{-k})$ denotes the number of good outcomes for the $K - 1$ other firms: in view of the symmetry assumption, the number $n = n(\boldsymbol{s}^{-k})$ is all that is needed to characterize the realizations $\boldsymbol{s}^{-k}$ of the other firms. To simplify notation let $\rho(e, \eta)$ denote the probability of a good outcome for a firm when its manager's effort is e and the common shock is η, i.e. $p_k(g, e_k, \eta) = \rho(e_k, \eta)$ and $p_k(b, e_k, \eta) = 1 - \rho(e_k, \eta)$, $k \in \mathcal{K}$. Using subscripts for partial derivatives, we assume $\rho_e > 0$, i.e. effort increases the probability of the good outcome, and $\rho_\eta > 0$, i.e. high values of η are favorable. $\rho_\eta > 0$ implies that (A2) holds. Since the derivatives of the likelihood function L for the good and the bad outcome are given by

[5] As a mathematical remark, note that the monotonicity properties of the optimal reward functions $\tau_k(s_k, \boldsymbol{s}^{-k})$ depend on log-super(sub)modularity of the underlying probability functions $p_k(s_k, e_k, \eta)$, $k \in \mathcal{K}$. Log-supermodularity in (s_k, e_k) implies monotonicity in s_k, and log-supermodularity in (s_k, η), combined with log-sub or supermodularity in (e_k, η) implies monotonicity in $\boldsymbol{s}^{-k}$.

$$L_\eta(g, e, \eta) = \frac{\rho_{e\eta}\rho - \rho_e\rho_\eta}{\rho^2}, \quad L_\eta(b, e, \eta) = \frac{-\rho_{e\eta}(1-\rho) - \rho_e\rho_\eta}{(1-\rho)^2}$$

the characteristics of the reward schedule $\tau_k(s_k, n(\boldsymbol{s}^{-k}))$ depend on the sign of the cross partial derivative $\rho_{e\eta}$.

(a) $\rho_{e\eta} = 0$

The likelihood function L is decreasing in η for both outcomes and the optimal reward schedule satisfies $\bar{\tau}_k(b, n) < \bar{\tau}_k(g, n)$ (since $\rho_e > 0$) and $\bar{\tau}_k(s_k, n) < \bar{\tau}_k(s_k, n')$ if $n > n'$. The reward schedule is "tournament-like" in that the more other agents there are who have a good outcome, the less manager k is paid.

(b) $\rho_{e\eta} \neq 0$

(i) $\rho_{e\eta} > 0$. The likelihood function is decreasing in η for the low outcome, $L_\eta(b, e, \eta) < 0$, but for the high outcome the sign is ambiguous. If ρ is given by $\rho(e, \eta) = a + b\, e^\alpha\, \eta^\beta$ with $a > 0,\ b > 0,\ a + b < 1,\ 0 < \alpha < 1,\ \beta > 0$ then $L_\eta(g, e, \eta) > 0$. In this case the reward $\bar{\tau}_k(b, n)$ decreases when n increases, while $\bar{\tau}_k(g, n)$ is an increasing function of n. When few other firms have good outcomes, η is likely to be low and effort is not likely to have much effect, so that a good or bad outcome for firm k has to be attributed to chance. When more firms have good outcomes, signaling a higher η, the managers's effort is more likely to have an effect so that it is worthwhile to reward the manager when the outcome is good and punish him when it is bad.

(ii) $\rho_{e\eta} < 0$. L is decreasing in η for the good outcome and has an ambiguous sign for the bad outcome. If ρ is given by $\rho(e, \eta) = a + b\,(e + \eta)^\alpha$ with $0 \leq e \leq 1/2,\ 0 \leq \eta \leq 1/2,\ a > 0,\ b > 0,\ a + b < 1,\ 0 < \alpha < 1$ and $(1 - \alpha) > b/(1 - a)$, then $L_\eta(b, e, \eta) > 0$. In this case $\bar{\tau}_k(g, n)$ decreases when n increases, while $\bar{\tau}_k(b, n)$ is an increasing function of n. Because of the decreasing returns property in $e + \eta$, a high value of η implies that the marginal effect of effort is low. Thus observing a high number of good outcomes for the other firms makes it unlikely that either a good or a bad outcome is the result of effort. As n decreases, the reward for a good outcome, and the punishment for a bad outcome, increase. Thus, while in case (i) the biggest differential between a good and a bad outcome for manager k occurs when many firms have good outcomes, in case (ii) it occurs when few firms have good outcomes.

For simplicity of exposition we have focused on the case where the outcome is a discrete random variable but it is clear that Proposition 1 applies to models in which the outcome is a continuous random variable, with density replacing probability mass in Assumption (A2) and in the definition of the local likelihood ratio.

Example 2 In examples of continuous outcomes models with a common shock studied in the literature, η enters either additively as in the model of Lazear and Rosen (1981) and Green and Stokey (1983) with $h^k(e_k, \epsilon_k, \eta) = z(e_k, \epsilon_k) + \eta$, or multiplicatively as in the model of Nalebuff and Stiglitz (1983) with $h^k(e_k, \epsilon_k, \eta) = e_k\eta + \epsilon_k$. In all cases $(\epsilon_1, \ldots, \epsilon_K)$ are i.i.d. and independent.

Let us show that the optimal reward schedule is tournament-like in the additive case while the reward can be either increasing or decreasing in the performance of others when the common shock affects the marginal product of effort.

(a) η does not affect the marginal product of effort

Let $h(e, \eta, \epsilon) = z(e, \epsilon) + \eta$ be the production function common to all firms where the distribution of z given e has a density $f(z, e)$. The density function for the output y given the manager's effort e and the common shock η is given by $\widetilde{f}(y, e, \eta) = f(y - \eta, e)$. In order that (A2) be satisfied the density f must be log-concave, i.e

$$\frac{f_z(z, e)}{f(z, e)} \text{ is decreasing in } z.$$

This is not a demanding assumption since most standard distributions (normal, gamma, chi square, Poisson, exponential) are log-concave, as well as many of the examples given by LiCalzi and Spaeter (2003). We assume that in addition the standard Monotone Likelihood Condition holds, i.e.

$$\frac{f_e(z, e)}{f(z, e)} \text{ is increasing in } z.$$

Then the local likelihood function $L(y, e, \eta) = f_e(y - \eta, e)/f(y - \eta, e)$ is a decreasing function of η: for a given realization of a firm, if η is higher, z is lower and, since MLRC holds, this tends to signal less effort on the part of the manager. Since (A2) is satisfied, for any realization y^k, the pay of manager k is a decreasing function of the outcomes of the other firms.

(b) η affects the marginal product of effort

Consider a more general version of the model of Nalebuff-Stiglitz where all firms have the production function $h(e, \eta, \epsilon) = \phi(e, \eta) + \epsilon$, with $\phi > 0$, $\phi_e > 0$, $\phi_\eta > 0$ where ϕ describes the production due to effort and the common shock η, and the idiosyncratic shock ϵ is additive. To ensure that (A2) holds we assume that the density of the idiosyncratic shock $f(\epsilon)$ is log-concave. The density function for the outcome y given e and η is $\widetilde{f}(y, e, \eta) = f(y - \phi(e, \eta))$ and the function L is given by $L(y, e, \eta) = -\phi_e(e, \eta) f'(y - \phi(e, \eta))/f(y - \phi(e, \eta))$. It is difficult to sign L_η without making more specific assumptions on the form of the density function f. The standard assumption is that the idiosyncratic shock is normally distributed with mean zero and variance σ^2. Then $L(y, e, \eta) = (1/\sigma^2)\phi_e(e, \eta)(y - \phi(e, \eta))$ and

$$L_\eta(y, e, \eta) = \frac{1}{\sigma^2}\left(\phi_{e\eta} y - (\phi_{e\eta}\phi + \phi_e\phi_\eta)\right).$$

(i) $\phi_{e\eta} > 0$. An increase in η increases the marginal product of effort. If $y < 0$ then $L_\eta(y, e, \eta) < 0$: when a low outcome is observed for firm k, the higher the realizations of other firms, the more likely it is that η was high and that effort was productive, and the more likely that the bad outcome can be attributed to shirking. When y is positive, $L_\eta(y, e, \eta)$ may not have the same

sign for all values of η, but the sign is positive for sufficiently high outcomes, provided ϕ is bounded. For example if $\phi(e,\eta) = e^{\alpha}\eta^{\beta}$, with $0 < \alpha < 1$, $\beta > 0$, $e \in [0, e^{max}]$, $\eta \in [0, \eta^{max}]$, then $\phi_{e\eta}\phi + \phi_e\phi_\eta = 2\phi\phi_{e\eta}$ and $L_\eta(y, e, \eta) = (1/\sigma^2)\phi_{e\eta}(y - 2\phi) > 0$ if $y > 2\phi(e^{max}, \eta^{max})$. This case is the analogue for the model with continuous outcomes of case b(i) in Example 1.
(ii) $\phi_{e\eta} < 0$. To sign $\phi_{e\eta}\phi + \phi_e\phi_\eta$, let us assume that $\phi(e,\eta) = (e+\eta)^{\alpha}$, with $0 < \alpha < 1$. If $0 < \alpha < 1/2$, $\phi_{e\eta}\phi + \phi_e\phi_\eta < 0$, so that if $y < 0$, then $L_\eta(y, e, \eta) > 0$. In this case the decreasing returns are very strong: a higher value of η decreases the productivity of effort so that a bad outcome is less likely to be due to lack of effort and the punishment decreases. For $y > 0$ the sign of L_η may not be constant but it is negative for high values of y ($y > (1-2\alpha)/(1-\alpha)\phi(e^{max}, \eta^{max})$) if ϕ is bounded. If $\alpha = 1/2$, $\phi_{e\eta}\phi + \phi_e\phi_\eta = 0$, so that $L_\eta > 0$ for $y < 0$ and $L_\eta < 0$ for $y > 0$. If $\alpha > 1/2$, $\phi_{e\eta}\phi + \phi_e\phi_\eta > 0$ so that when $y > 0$, $L_\eta < 0$. For $y < 0$ the sign may not be constant but is positive for low values of y provided ϕ is bounded. The case $\phi_{e\eta} < 0$ is thus the analogue of case b(ii) in Example 1.

5 Conclusion

The discussion of relative performance compensation of CEOs in corporate finance generally uses the simplest additive model ($h^k(e_k, \epsilon_k, \eta) = e_k + \epsilon_k + \eta$) as the reference model (see e.g. Gibbons and Murphy 1990). It is argued that relative performance evaluation is valuable because it factors out the effect of common shocks—it avoids exposing CEOs to risks which do not serve to create incentives and for which they would otherwise need to be compensated. Relative performance evaluation implies that a CEO's compensation should be a decreasing function of the outcomes of other firms. Murphy (1999), however, reports that only 20% of large US companies explicitly use relative performance criteria to determine CEO compensation. On the other hand the same survey shows that the majority of large corporations use stock options and that in the last 10 years they have become the most significant component of CEO compensation. Although stock options could be indexed to the market—to make them adhere to the relative performance criterion—in practice they are not. As a result, the compensation of a CEO is higher when the overall level of economic activity and the stock market are higher.

From the above analysis this type of compensation may be justified if the general state of the economy has a positive effect on the productivity of the top executive. Indeed it seems plausible, when entrepreneurship and innovation are the qualities required, that the actions of a CEO will have their greatest impact in good times, when the economy is expanding and has the greatest capacity to absorb new products or new technologies. However, if the main contribution of the CEO consists in steering the firm through difficult times, then the compensation should be higher when the firm does well while the market as a whole is depressed, and in this case stock options are not an appropriate type of compensation. Thus it seems that a model like that in Example 2, which specifies how the economic environment affects the productivity of managerial input, may be useful for assessing whether CEO compensation should, or should not, factor out industry and economic trends.

Appendix

Proof of Lemma 1 If the pair $(\bar{\boldsymbol{\tau}}_k, \bar{e}_k)$ is optimal for firm k then $\bar{\boldsymbol{\tau}}_k$ minimizes the expected cost of inducing the effort $\bar{e}_k$. Let $\bar{C}_k(\boldsymbol{\tau}_k) = \sum_{\boldsymbol{s}\in\mathcal{S}} P(\boldsymbol{s}, \bar{\boldsymbol{e}})\boldsymbol{\tau}_k(\boldsymbol{s})$ denote the expected cost of $\boldsymbol{\tau}_k$ when the probability is $(P(\boldsymbol{s}, \bar{\boldsymbol{e}}))_{\boldsymbol{s}\in\mathcal{S}}$ and let

$$\bar{U}_k(\boldsymbol{\tau}_k, e_k) = \sum_{\boldsymbol{s}\in\mathcal{S}} P\left(\boldsymbol{s}, e_k, \bar{\boldsymbol{e}}^{-k}\right) u_k(\tau_k(\boldsymbol{s})) - c_k(e_k)$$

denote the expected utility of manager k, net of the cost of effort, when the other managers make effort $\bar{\boldsymbol{e}}^{-k}$. Then $\bar{\boldsymbol{\tau}}_k$ is solution of the problem

$$\begin{aligned} &\min \bar{C}_k(\boldsymbol{\tau}_k) \\ \text{subject to: } &\bar{U}_k(\boldsymbol{\tau}_k, \bar{e}_k) \geq \bar{v}_k \\ &\bar{e}_k \in \arg\max \bar{U}_k(\boldsymbol{\tau}_k, e_k) \end{aligned} \tag{6}$$

The optimality of $\bar{e}_k$ for manager k implies $D_{e_k}\bar{U}_k(\bar{\boldsymbol{\tau}}_k, \bar{e}_k) = 0$. Consider a marginal change $\bar{\boldsymbol{\tau}}_k \to \bar{\boldsymbol{\tau}}_k + d\boldsymbol{\tau}_k$ such that $D_{e_k}\bar{U}_k(\bar{\boldsymbol{\tau}}_k + d\boldsymbol{\tau}_k, \bar{e}_k) = 0 \Longleftrightarrow D_{e_k,\boldsymbol{\tau}_k}\bar{U}_k(\bar{\boldsymbol{\tau}}_k, \bar{e}_k)\cdot d\boldsymbol{\tau}_k = 0$. (A1) implies that $\bar{e}_k$ remains optimal for $\bar{\boldsymbol{\tau}}_k + d\boldsymbol{\tau}_k$. Suppose not. Then there exists $\tilde{e}_k \neq \bar{e}_k$ such that $\bar{U}_k(\bar{\boldsymbol{\tau}}_k + d\boldsymbol{\tau}_k, \tilde{e}_k) > \bar{U}_k(\bar{\boldsymbol{\tau}}_k + d\boldsymbol{\tau}_k, \bar{e}_k)$ and, by continuity of $\bar{U}_k$, $\bar{U}_k(\bar{\boldsymbol{\tau}}_k, \tilde{e}_k) \geq \bar{U}_k(\bar{\boldsymbol{\tau}}_k, \bar{e}_k)$, contradicting the uniqueness of the maximum. Thus every $d\boldsymbol{\tau}_k \in \Re^S$ satisfying

$$\begin{aligned} D_{\boldsymbol{\tau}_k}\bar{U}_k(\bar{\boldsymbol{\tau}}_k, \bar{e}_k) \cdot d\boldsymbol{\tau}_k &\geq 0 \\ D_{e_k,\boldsymbol{\tau}_k}\bar{U}_k(\bar{\boldsymbol{\tau}}_k, \bar{e}_k) \cdot d\boldsymbol{\tau}_k &= 0 \end{aligned}$$

[the local version of the constraints in (6)] must satisfy $D_{\boldsymbol{\tau}_k}\bar{C}_k(\bar{\boldsymbol{\tau}}_k) \cdot d\boldsymbol{\tau}_k \geq 0$. By the Minkowski-Farkas Lemma there exist $(\lambda_k, \mu_k) \in \Re^2$ such that

$$D_{\boldsymbol{\tau}_k}\bar{C}_k(\bar{\boldsymbol{\tau}}_k) = \lambda_k D_{\boldsymbol{\tau}_k}\bar{U}_k(\bar{\boldsymbol{\tau}}_k, \bar{e}_k) + \mu_k D_{e_k,\boldsymbol{\tau}_k}\bar{U}_k(\bar{\boldsymbol{\tau}}_k, \bar{e}_k)$$

which is equivalent to (3). Note that (3) implies, by summing over $\boldsymbol{s}$

$$\sum_{\boldsymbol{s}\in\mathcal{S}} P(\boldsymbol{s}, \bar{\boldsymbol{e}})\left(\frac{1}{u_k'(\bar{\tau}_k(\boldsymbol{s}))} - \lambda_k\right) = \sum_{\boldsymbol{s}\in\mathcal{S}} \frac{\partial P(\boldsymbol{s}, \bar{\boldsymbol{e}})}{\partial e_k} = 0 \Longrightarrow \lambda_k = E\left(\frac{1}{u_k'(\bar{\boldsymbol{\tau}}_k)}\right) > 0$$

and (3) and $D_{e_k}\bar{U}_k(\bar{\boldsymbol{\tau}}_k, \bar{e}_k) = 0$ imply

$$\begin{aligned} \mu_k c'(\bar{e}_k) &= E\left(\frac{1}{u_k'(\bar{\boldsymbol{\tau}}_k)} \cdot u_k(\bar{\boldsymbol{\tau}}_k)\right) - E\left(\frac{1}{u_k'(\bar{\boldsymbol{\tau}}_k)}\right) E\left(u_k(\bar{\boldsymbol{\tau}}_k)\right) \\ &= \operatorname{cov}\left(\frac{1}{u_k'(\bar{\boldsymbol{\tau}}_k)}, u_k(\bar{\boldsymbol{\tau}}_k)\right) > 0. \end{aligned}$$

The positive sign for the covariance comes from the fact that, by concavity of u_k, $1/(u_k'(\tau_k(\boldsymbol{s})))$ and $u_k(\bar{\boldsymbol{\tau}}_k)$ are positively dependent random variables,[6] with $\operatorname{var}(\bar{\boldsymbol{\tau}}_k) > 0$ since the optimal effort is interior. Thus $\lambda_k > 0$, $\mu_k > 0$.

[6] The covariance of positively dependent random variables is positive, see Magill and Quinzii (1996, p. 170).

References

Celentani, M., Loveira-Pazo, R.: What form of relative performance evaluation?, Universitat Pompeu Fabra, Economics Working Papers (2004)
Gibbons, R., Murphy, K.J.: Relative performance evaluation for chief executive officers. Ind Labor Relations Rev **43**, 30–51 (1990)
Green, J., Stokey, N.: A comparison of tournaments and contracts. J Pol Econ **91**, 349–364 (1983)
Grossman, S., Hart, O.: An analysis of the principal-agent problem. Econometrica **51**, 7–45 (1983)
Himmelberg, C.P., Hubbard, R.G.: Incentive pay and the market for CEOs: an analysis of pay-for-performance sensitivity. Columbia University Discussion Paper (2000)
Holmström, B.: Moral hazard in teams. Bell J Econ. **13**, 324–340 (1982)
Lazear, E., Rosen, S.: Rank-order tournaments as optimum labor contracts. J Pol Econ **89**, 841–864 (1981)
LiCalzi, M., Spaeter, S.: Distributions for the first-order approach to principal agent problems. Econ Theory **21**, 167–173 (2003)
Magill, M., Quinzii, M.: The theory of incomplete markets. Cambridge, Massachusetts: MIT Press, 1996
Milgrom, P.: Good news and bad news: representation theorems and applications. Rand J Econ **12**, 380–391 (1981)
Mookherjee, D.: Optimal incentive schemes with many agents. Rev Econ Stud **51**, 433–446 (1984)
Murphy, K.J.: Executive compensation. In: Ashenfelter, O., Card, O.(eds.) Handbook of labor economics, Vol. 3B, Chap. 38, pp. 2485–2563. Amsterdam: North-Holland 1999
Nalebuff, B.J., Stiglitz, J.E.: Prizes and incentives: towards a general theory of compensation and competition, Bell Econ **14**, 21–43 (1983)
Rogerson, W.P.: The first-order approach to principal-agent problems. Econometrica **53**, 1357–1367 (1985)
Sinclair-Desgagné, B.: The first-order approach to multi-signal principal-agent problems. Econometrica **62**, 459–465 (1994)