COMMON SHOCKS AND RELATIVE COMPENSATION SCHEMES

Michael MAGILL

Department of Economics
University of Southern California
Los Angeles, CA 90089-0253
magill@usc.edu

Martine QUINZII

Department of Economics
University of California
Davis, CA 95616-8578
mmquinzii@ucdavis.edu

December 2004

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 derive conditions on the way the common shock affects the marginal product of effort under which the likelihood ratio is decreasing for all output levels, or increasing for some output levels and decreasing for others.

Common Shocks and Relative Compensation Schemes

1. Introduction

It has been shown by Holsmtröm (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. Holsmtröm also showed that under specific assumptions on the production function and under normality assumptions 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. However little has been established on the way an optimal contract makes use of the information provided by the realized performance of the other agents. Should the reward of an agent decrease or increase when the performance of agents in the comparison group increases?

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. 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 to occur 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.

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-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-Stiglitz (1983), a higher shock tends to raise the productivity of effort. Then 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. In the conclusion, we discuss the application of these results to executive compensation.

2. Model

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 k can only manage firm k or take an outside option which determines his reservation utility of 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. Thus $y_k = h_k(e_k, \eta, \epsilon_k)$, where $e_k \in \Re_+$, $k \in K$.[1] We assume that h_k can only take a countable number of values[2] indexed in increasing order by $s_k \in S_k = \{1, \cdots, S_k\}$: that is $s_k > s'_k \implies y_{s_k} > y_{s'_k}$. The idiosyncratic and common shocks $\{\epsilon_1, \cdots, \epsilon_K, \eta\}$ are assumed to be unobservable independent random variables. For given e_k and η, the distribution function of ϵ_k induces a probability distribution $p_k(\cdot, e_k, \eta)$ on S_k whose cumulative distribution function is denoted by F_k. That is, $F_k(\alpha, e_k, \eta) \doteq \sum_{\{s_k | y^k_{s_k} \leq \alpha\}} p_k(s_k, e_k, \eta)$. For all $k \in K$ the probabilities $(p_k)_{k \in K}$ are assumed to have the following properties:

(A1) $p_k(s_k, e_k, \eta) > 0$ for all $(s_k, e_k, \eta) \in S_k \times \Re_+ \times \Re$ and p_k is a differentiable function of e_k.

(A2) For all $e_k > 0$, and $\eta \in \Re$, $\dfrac{\frac{\partial}{\partial e_k} p_k(s_k, e_k, \eta)}{p_k(s_k, e_k, \eta)}$ is an increasing function of s_k.

(A3) For all $\eta \in \Re$, and $\min_{s_k}(y^k_{s_k}) \leq \alpha < \max_{s_k}(y^k_{s_k})$, $1 - F_k(\alpha, e_k, \eta) \doteq \sum_{\{s_k | y^k_{s_k} > \alpha\}} p_k(s_k, e_k, \eta)$ is a concave increasing function of e_k.

(A4) For any $e_k \in \Re+$, $p_k(s_k, e_k, \eta)$ is log-supermodular in (s_k, η), i.e. if $\eta > \eta'$ the ratio $p_k(s_k, e_k, \eta)/p_k(s_k, e_k, \eta')$ is increasing in s_k.

[1] We use the same notation for a set and for the number of its elements.

[2] The proof of the theorem carries over without change to the case of continuous outcomes provided that we can use the "first-order approach". As Jewitt pointed out, Assumption A3 may not be satisfied in a model with continuous outcomes and standard assumptions on the shocks, in which case it may be replaced by the assumptions suggested by Jewitt (1988). To simplify the exposition we present the model and the assumptions needed for the first-order approach to be valid only in the countable case (see Rogerson (1985)).

Thus for a given value of the common shock η, $p_k(s_k, e_k, \eta)$ satisfies the standard assumptions: by (A2) it satisfies the Monotone Likelihood Ratio Condition (MLRC) according to which if $e_k > e'_k$ the likelihood ratio $\frac{p_k(s_k, e_k, \eta)}{p_k(s_k, e'_k, \eta)}$ is increasing in s_k (Milgrom (1981)), i.e p_k is log-supermodular in (s_k, e_k). This implies that a larger effort leads to a stochastically dominant shift in F_k, but (A3) adds the condition of stochastic decreasing returns to effort. (A4) is the condition needed to ensure that a high realization of firm k can be interpreted as a signal that the common shock η has been favorable: for the same level of effort, a higher value of η implies a greater likelihood of high outcomes, or, since the property of log-supermodularity is symmetric, observing a higher production for a firm increases the likelihood that the shock η has been favorable. A4 implies that if $\eta > \eta'$ the distribution function $F_k(\alpha, e_k, \eta)$ stochastically dominates $F_k(\alpha, e_k, \eta')$ (see Rogerson (1985)), and, since this is true for all firms, high values of η constitute a positive shock, while low values of η are negative shocks for the economy.

Let $\boldsymbol{S} = S_1 \times \ldots \times S_K$. A state of the economy is a realization $\boldsymbol{s} = (s_1, \ldots, s_K) \in \boldsymbol{S}$, 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. Since the idiosyncratic shocks are independent, for a given vector $\boldsymbol{e}$ and a given value of the external shock η the probability of observing state $\boldsymbol{s}$ is

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

Let $G(\eta)$ be the distribution function of η. When the realization of η is not known the probability of $\boldsymbol{s}$ given $\boldsymbol{e}$ is

$$P(\boldsymbol{s}, \boldsymbol{e}) \doteq \int_{\Re} p(\boldsymbol{s}, \boldsymbol{e}, \eta) dG(\eta)$$

since η is independent of the idiosyncratic shocks. We assume that the random variables $\{\epsilon_1, \cdots, \epsilon_K, \eta\}$ are not observable by any agent but that the structure of the economy $(p_1, \ldots, p_K, G)$ is common knowledge.

Each firm k is owned by a collection of risk-neutral shareholders, who hire the manager and offer an incentive contract τ^k which guarantees the manager's reservation utility level. Each manager is risk averse and, given the pay schedule τ^k and effort e_k has utility

$$U_k(\tau^k, e_k) = E(u_k(\tau_k)) - c_k(e_k)$$

where $u_k(\cdot), c_k(\cdot)$ are differentiable, increasing, $u_k(\cdot)$ is concave and $c_k(\cdot)$ is convex. Let $\bar{\nu}_k$ denote the manager's reservation utility.

Because the realization $\boldsymbol{s}^{-k}$ of firms other than k contains information about the common shock η, the optimal contract τ^k for manager k will use this information to provide incentives at least cost for the shareholders. The contract τ^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. We will restrict attention to interior Nash equilibria where all managers are induced to make a positive effort.

Definition. $(\bar{\boldsymbol{\tau}}, \bar{\boldsymbol{e}}) = (\bar{\tau}^1, \ldots, \bar{\tau}^K, \bar{e}_1, \ldots, \bar{e}_K)$ with $\bar{\boldsymbol{e}} \gg 0$ is an interior *Nash equilibrium with optimal incentive contracts* if for each $k \in K$, $(\bar{\tau}^k, \bar{e}_k)$ solves the problem $(\mathcal{P}_k)$

$$\max_{(\tau_k, e_k) \in \Re^S \times \Re_+} \sum_{s \in S} P(\boldsymbol{s}, e_k, \bar{\boldsymbol{e}}^{-k})(y^k_{s_k} - \tau^k(\boldsymbol{s}))$$

subject to

$$\sum_{s \in S} P(\boldsymbol{s}, e_k, \bar{\boldsymbol{e}}^{-k}) u_k(\tau^k(\boldsymbol{s})) - c_k(e_k) \geq \bar{\nu}_k \tag{PC$_k$}$$

$$\sum_{s \in S} \frac{\partial}{\partial e_k} P(\boldsymbol{s}, e_k, \bar{\boldsymbol{e}}^{-k}) u_k(\tau^k(\boldsymbol{s})) - c'_k(e_k) = 0 \tag{IC$_k$}$$

Remark. It is easy to verify that under A1-A3, concavity of u_k and convexity of c_k, the first-order condition (IC$_k$) characterizes the optimal effort $\bar{e}_k > 0$ of manager k (even though the contract depends on $\boldsymbol{s}^{-k}$), and that the associated multiplier is positive.

3. Result

The properties of the optimal contract $\bar{\tau}^k$ at a Nash equilibrium $(\bar{\boldsymbol{\tau}}, \bar{\boldsymbol{e}})$ can be derived from the first-order conditions for the maximum problem $(\mathcal{P}_k)$ with respect to $\tau^k = (\tau^k_s)_{s \in S}$ which are

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

where $(\lambda_k, \mu_k) \gg 0$ are the multipliers associated with the participation and incentive constraints (PC$_k$) and (IC$_k$). Define the *local likelihood function* $L_k : S_k \times \Re_+ \times \Re \to \Re$ by

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

Proposition. *Let (A1)-(A4) be satisfied. For any realization $s_k \in S_k$, the optimal reward schedule $\bar{\tau}^k(s_k, \boldsymbol{s}^{-k})$ 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.*[3]

Proof: ($\Leftarrow$) Suppose $L_k(s_k, \bar{e}_k, \eta)$ is decreasing in η. We want to show that if $s, s' \in S$ 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 $\mu_k > 0$ and u'_k is strictly decreasing it follows from (F_τ) that what we need to show is that $A < 0$, where A is defined by

$$A \doteq \frac{\frac{\partial}{\partial e_k} P(\boldsymbol{s}, \bar{\boldsymbol{e}})}{P(\boldsymbol{s}, \bar{\boldsymbol{e}})} - \frac{\frac{\partial}{\partial e_k} P(\boldsymbol{s}', \bar{\boldsymbol{e}})}{P(\boldsymbol{s}', \bar{\boldsymbol{e}})} \tag{1}$$

Note that

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

where

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

Note that for all $\boldsymbol{s} \in \boldsymbol{S}$, $a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta) > 0$, $\int_{\Re} a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta) dG(\eta) = 1$, and

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

By (A4), 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 $\lambda(\eta) \doteq a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta)/a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \eta)$ is an increasing function of η.[4] Since $\int_{\Re} a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta) dG(\eta) = \int_{\Re} \lambda(\eta) a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \eta) dG(\eta) = \int_{\Re} a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \eta) dG(\eta) = 1$, $\lambda(\eta)$ cannot be always 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 $\int_{\eta \leq \bar{\eta}} dG(\eta) > 0$, $\int_{\eta > \bar{\eta}} dG(\eta) > 0$.

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

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

$$L_k(s_k, \bar{e}_k, \eta) \left(a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta) - a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \eta) \right) \leq L_k(s_k, \bar{e}_k, \bar{\eta}) \left(a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta) - a(\boldsymbol{s}', \bar{\boldsymbol{e}}, \eta) \right) \tag{2}$$

[3] i.e. if $p_k(s_k, \bar{e}_k, \eta)$ is log-submodular (log-supermodular) in (e_k, η).

[4] For sake of completeness and to tie this part of the proof with the "only if" part, we give the direct proof without using the fact that a is supermodular in $(\boldsymbol{s}, \eta)$ and that this in turn implies a property of first-order stochastic dominance.

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 (2) is satisfied with a strict inequality. Thus

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

If the function $L_k(s_k, \bar{e}_k, \cdot)$ is increasing in η then inequality (2) 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 $\int_{\Re} a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta) dG(\eta) = \int_{\Re} a(\boldsymbol{s'}, \bar{\boldsymbol{e}}, \eta) dG(\eta) = 1$, it must be that $a(s, \bar{e}, \eta') - a(s', \bar{e}, \eta') < 0$ and $a(\boldsymbol{s}, \bar{\boldsymbol{e}}, \eta) - a(\boldsymbol{s'}, \bar{\boldsymbol{e}}, \eta) > 0$. Thus A defined in (1) is such that

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

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. $\square$

Remark. To create incentives at minimum cost, the optimal contract $\bar{\tau}^k$ must reward the manager in circumstances which are most likely to occur when the agent makes a high rather than a low effort and punish the agent in circumstances which are more likely with a low effort. If e_k is a high effort level and e'_k a lower effort level, the relative likelihood of observing s_k when the effort is e_k rather than e'_k is

$$\frac{p_k(s_k, e_k, \eta)}{p_k(s_k, e'_k, \eta)} = \exp \int_{e'_k}^{e_k} L_k(s_k, t, \eta) dt$$

If the local likelihood L_k at s_k decreases when η increases, then the relative likelihood that s_k is observed with e_k rather than e'_k decreases (or the relative likelihood that s_k is observed with e'_k rather than e_k increases). If the shareholders could observe the common shock and base the contract on η, the reward of the manager would decrease when η increases. When η is not observable, the realizations $\boldsymbol{s}^{-k}$ of the firms other than k give information on the value of η: since by A4 the likelihood of high outcomes increase with η, higher values for $\boldsymbol{s}^{-k}$ lead to a higher estimate of η and a lower reward for manager k.

4. Examples

We give two examples of settings where the Proposition can be used to analyze properties of the optimal reward schedule.

Example 1. Consider the simple 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, 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 aggregate shock is η, i.e. $p_k(g, e_k, \eta) = \rho(e_k, \eta)$ and $p_k(b, e_k, \eta) = 1 - \rho(e_k, \eta)$, $k \in K$. Using subscripts for partial derivatives, A1-A4 are satisfied if $\rho_e > 0$, $\rho_\eta > 0, \rho_{ee} \leq 0$. Since the derivatives of the likelihood function L for the good and the bad outcome are given by

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

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 $\bar{\tau}^k(b, n) < \bar{\tau}^k(g, n)$ (because of A2) 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 when the outcome is good and punish 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 $0 \leq e \leq 1/2,\ 0 \leq \eta \leq 1/2,\ a > 0,\ b > 0,\ a + b \leq 1,\ 0 < \alpha < 1$ and $(1-\alpha)/\alpha > b/(1-a)$, 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 the proposition applies to models in which the outcome is a continuous random variable (with density replacing probability mass), provided A4 is satisfied and the optimal contract satisfies the FOCs (F_τ) with $\mu_k > 0$.

Example 2. In standard continuous outcomes models with a common shock, η enters either additively as in the model of Lazear-Rosen (1981) and Green-Stokey (1983) with $h_k(e_k, \epsilon_k, \eta) = z(e_k, \epsilon_k) + \eta$, or multiplicatively as in the model of Nalebuff-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 of η.

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 $f(z, e)$ which is log-concave and satisfies MLRC, i.e.

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

f is assumed to be log-concave to ensure that A4 holds: this is not a demanding assumption since most standard distributions (normal, gamma, chi square, Poisson, exponential, and more) are log-concave. The density function of the output y given the manager's effort e and the aggregate shock

η is given by $\widetilde{f}(y, e, \eta) = f(y - \eta, e)$, with the local likelihood function $L(y, e, \eta) = f_e(y - \eta, e)/f(y - \eta, e)$. If f satisfies MLRC, then L 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 A4 is satisfied, for any realization y_k, the pay of manager k is a decreasing function of the outcomes of the other agents.

(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 aggregate shock η, and the idiosyncratic shock ϵ is additive. To ensure that A4 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 ($y > \frac{1-2\alpha}{1-\alpha}\phi(e^{max}, \eta^{max})$) 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-Murphy (1990)). It is argued that relative performance evaluation is valuable because it factors out the effect of agggregate shocks and eliminates unnecessary risks from the compensation: relative performance evaluation implies that a CEO's compensation should be a decreasing function of the outcomes of other firms. Murphy's survey (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 ten years they have become the most significant component of CEO compensation. Although stock options could be indexed on 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 positively affects the productivity of the top executive. And it seems plausible, when entrepreuneurship 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 principal part of the CEO's contribution is to steer 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(b), which obliges us to specify 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.

References

Gibbons, R. and Murphy, K.J.(1990), "Relative Performance Evaluation for Chief Executive Officers", *Industrial and Labor Relations Review*, 43, 30-51.

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

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

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

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

Milgrom, P. (1981), "Good News and Bad News: Representation Theorems and Applications", *Rand Journal of Economics*, 12, 380-391.

Mookherjee, D. (1984), "Optimal Incentive Schemes with Many Agents", *Review of Economic Studies*, 51, 433-446.

Murphy, K.J.(1999), "Executive Compensation", *Handbook of Labor Economics, Volume 3B*, O. Ashenfelter and D. Card eds, North-Holland.

Nalebuff, B.J. and J.E. Stiglitz (1983), "Prizes and Incentives: Towards a General Theory of Compensation and Competition", *Bell Journal of Economics*, 14, 21-43.

Rogerson, W.P., "The First Order Approach to Principal-Agent Problems", *Econometrica*, 1985, 53, 1357-1367.