ANNALES D'ÉCONOMIE ET DE STATISTIQUE. – N° 25/26 — 1992

Biased Contests and Moral Hazard: Implications for Career Profiles

Margaret A. MEYER *

ABSTRACT. – We study the design of a sequence of two contests between a pair of identical risk averse employees whose effort choices are private information. It is optimal for the organization to "bias" the second contest in favor of the early winner – the reduction in second-period incentives is outweighed by the increase in first-period incentives. Thus, even though first-period success reflects only transitory shocks and not ability, it is efficient to structure the contests so these shocks have persistent effects on employees' careers.

Biais dans les compétitions et aléa moral : implications pour les profils de carrières

RÉSUMÉ. – On examine une compétition formée de deux concours successifs entre deux employés identiques présentant de l'aversion à l'égard du risque et dont les choix d'efforts sont des informations privées. Il est optimal pour l'organisateur de la compétition de « biaiser » le second concours en faveur du premier, la réduction d'incitations dans la seconde étape étant plus que compensée par l'accroissement des incitations dans la première. Ainsi, même si le résultat du premier concours ne reflète que des chocs transitoires, aucunement des compétences, il est efficace d'organiser la compétition de sorte que les chocs affectent de manière permanente les carrières des employés.

---

* M. A. MEYER: Nuffield College, Oxford OX1 1NF, England. I am very grateful to Paul Klemperer and Dilip Mookherjee for valuable discussions and to seminar participants at Oxford and Stanford and an anonymous referee for helpful comments. Financial support and hospitality from the Stanford Institute for Theoretical Economics is gratefully acknowledged.

1 Introduction

This paper studies the implications, for the structure of jobs in organizations and the career profiles of employees, of limitations on the information available about employee performance. Organizational sociologists have found that earnings and promotions in the later stages of employees' careers are strongly correlated with earnings and promotions in the early stages (Kanter [1977], Rosenbaum [1984]). Later success is positively associated with early success, even when one controls for the effect of observable characteristics likely to affect performance, such as education. One potential source of this correlation is differences in ability that persist over time and that are not captured by the observable covariates. The companion paper, Meyer [1991], develops a simple model of an organization designing a sequence of contests in order to learn about differences in ability, and shows that the optimal learning strategy can reinforce the effect of these differences: early success can be even more strongly associated with later success under the optimal learning strategy than under a naive process of information accumulation.

Here we show that even in situations where employees are known to be of identical ability, an organization may induce a correlation between early and later success in order to limit the costs of moral hazard. In designing a sequence of two contests for a pair of workers, it is optimal to reward success in the first with an increased probability of success in the second. This increase can be accomplished by assigning early winners to more productive jobs or by providing them with extra training – treatment which puts them on the "fast track". Although the asymmetric treatment of early winners and losers reduces incentives for effort in the second contest, this cost is outweighed, with risk averse employees, by the increase in incentives in the first contest due to the future rewards for current success. Thus, even though the difference in employees' performance in the early contest reflects only transitory random factors and not ability, it is efficient for the organization to structure the sequence of contests so the transitory shocks have persistent effects on employees' careers.

The model is intended to represent organizations which keep new employees in junior positions for several periods before making major job assignment decisions. Before the major decision, there must be at least one opportunity for an interim evalution of employees' performance and for the subsequent treatment of employees to be differentiated according to its results. The employees may, for example, be junior lawyers, academics, accountants, or members of the military, but the analysis is not confined to organizations employing "up-or-out" policies. After working in positions explicitly designated for new workers, some employees may be moved "sideways", while others are moved up.

In the formal model, two risk averse agents compete against each other for two periods; the organization's major promotion decision is made after the second, and the interim evaluation after the first. The agents are productively identical but make privately observed choices of effort, so the organization (henceforth referred to as the principal) must design its job structures and compensation policies to provide incentives for effort.

We focus on situations in which the costs associated with gathering or transmitting information result in the principal obtaining only rank-order information about the agents' outputs in each period. However, we assume that the principal can costlessly choose the criterion, called the "bias", that determines how the outputs are ranked: if the chosen bias is c in agent i's favor, then agent i is declared the winner as long as his output does not fall short of his rival's by more than c.

When the "observer" who monitors the agents' performance obtains cardinal information but, because of the costs of communication, reports only ordinal information to the principal, the principal can implement bias by instructing the observer to use asymmetric evaluation criteria. But even if the observer can identify only which of two output levels is larger, the principal can control the level of bias by differentiating agents' tasks or work environments, or by providing different amounts of training or equipment. Examples of the latter form of bias are the assignment of employees to clients with different needs or attitudes and the provision by senior colleagues of different amounts of guidance.

With a production technology in each period in which output is additively separable in effort and the individual-specific component of the random shock, we show that the Nash equilibrium effort levels of the agents are equal, whatever the bias, and are decreasing in the magnitude of the bias. It follows that the principal will never introduce bias in the first contest, since doing so would reduce the agents' first-period incentives, as well as provide ex ante rents to the favored agent. Furthermore, a commitment by the principal to bias the second contest in favor of the first-period loser will not be advantageous because, relative to no bias, it would reduce incentives in both the first period (by creating a future punishment for current success) and the second period.

Our major result is that the principal's cost of implementing any given first- and second-period effort levels is minimized by committing to favor the first-period winner by a strictly positive amount in the second period. This result hinges on the fact that, starting with no bias, the introduction of a small amount in favor of the first-period winner generates a first-order increase in first-period incentives, but only a second-order reduction in second-period incentives; in consequence, the original effort levels can be induced with the imposition of less risk on the agents and hence at lower cost to the principal. It is thus optimal for the principal to commit to spreading rewards over time for risk averse agents, by designing jobs so that early success is rewarded by an increase in future promotion prospects as well as by a higher current salary. A correlation between early success and later success therefore results, even though agents have identical abilities and choose identical effort levels, so in equilibrium, success in the first period is due entirely to transitory random shocks.

Related Work

The companion paper, MEYER [1991], focuses on how biasing contests improves the information about relative abilities conveyed by rank-order observations on outputs. In that paper, the organization's payoff is assumed to depend only on the quality of the promotion decision that follows the observation periods; the incentive effects of bias are ignored. The key result is that the value of the final observation is maximized when the organization biases the final contest in favor of the employee with the better performance record. With respect to promotion chances, the optimal bias thus reinforces the advantage this employee derives from his likely edge in ability. In this paper, we abstract from the learning motive for bias by considering agents of identical ability, and we focus on the incentive costs and benefits, which accrue during the observation periods. Since in this setting, the benefits of bias accrue before its costs, these benefits can be realized only if the principal can commit to the use of bias in later contests, for example, through the design of jobs or training procedures.

Several other papers analyze settings in which it is optimal for an organization or principal to commit to selection criteria that, *ex post*, discriminate to an inefficiently large extent in favor of those who achieve early success, because such commitments bring incentive benefits.

LAFFONT and TIROLE [1988] study the principal's design of an incentive contract for an incumbent supplier, when a potential replacement will be available in the second period. When the first-period investment undertaken by the incumbent is unobservable but transferable to the entrant if the incumbent is replaced, then it is optimal for the principal to commit to favor the incumbent in choosing the second-period supplier. Introducing a small bias causes a first-order improvement in the incumbent's investment incentives but only a second-order loss from inefficient selection.

MILGROM and ROBERTS [1988] model employees who divert effort from current production to credential-building activities in order to influence promotion decisions. Even if current performance (in contrast to credentials) is uninformative about suitability for promotion, it is optimal for the organization to supplement the incentives for productive effort provided by performance bonuses by committing to base promotion decisions, with strictly positive probability, on current performance.

PRENDERGAST [1989] shows that when a firm acquires private information about workers' abilities, it may have an incentive to signal this information to workers, to induce high-ability workers to invest in firm-specific human capital. One type of credible signal of high ability is promotion to a job beyond the worker's capacity. The firm may thus choose to publicly distinguish among its workers at an early stage, putting some in "fast track" jobs for which they are insufficiently qualified.

Another setting in which relatively similar workers may have very different career paths is when "up-or-out" rules are employed, under which junior workers are dismissed if they are not promoted by a certain time. Despite the inefficiency of dismissing workers who could be profitably retained in junior positions, KAHN and HUBERMAN [1988], GILSON and

MNOOKIN [1989], and WALDMAN [1990] argue that commitment to such rules can be advantageous because it deters the organization from exploiting, by not promoting, junior workers whom it has (privately) identified as potentially productive senior workers; this deterrent in turn gives junior workers incentives to invest in human capital.

Section 2 of this paper describes the model of a sequence of two contests between identical employees, and Section 3 analyzes the consequences for second-period efforts of the use of bias. In Section 4, we show that an optimal two-period contract commits the principal to use second-period bias to favor the first-period winner. Section 5 extends the result on the value of committing to conduct an unfair contest to two simple settings in which the employees are not identical.

2 The Model

An organization, referred to as the principal, hires a pair of identical agents (labeled i and j) and employs them in entry-level positions for two periods. After the two periods, one of the agents is promoted to a higher level in the hierarchy. In each period t, agent k's output, x_k^t, is given by:

$$x_k^t = f^t(a_k^t, s^t) + \varepsilon_k^t, \qquad k = i, j, \qquad t = 1, 2,$$

where a_k^t is k's effort in period t, s^t is a common shock affecting the production of both agents in period t, and ε_k^t is the individual-specific shock in period t. In each period, the agents choose their efforts simultaneously and noncooperatively, and an agent's choice of effort is not observable by anyone else.

We assume that $f_1^t > 0$ and $f_{11}^t \leqslant 0$, where numerical subscripts denote partial derivatives. We define $\Delta\varepsilon^t \equiv \varepsilon_i^t - \varepsilon_j^t$ and assume that $\Delta\varepsilon^1$ and $\Delta\varepsilon^2$ are distributed independently and symmetrically about 0. Symmetry about 0 ensures that within each period there is no systematic difference between the agents' tasks, and independence ensures that any shock to relative outputs does not persist over time. Let $\Delta\varepsilon^t$ have support $(-\infty, \infty)$ and let $G^t(.)$ denote its cumulative distribution function and $g^t(.)$ its density. Assume that $g^t(.)$ is unimodal and is continuously differentiable and strictly positive everywhere. It follows that $g^t(.)$ is maximized at 0 and $g^{t\prime}(0) = 0$. Assume that s^1 and s^2 are distributed independently of each other and of $\Delta\varepsilon^1$ and $\Delta\varepsilon^2$.

The principal is risk neutral and can borrow and save at an interest rate corresponding to a discount factor $\delta_P \in (0, 1]$. He maximizes the discounted sum of expected profit over the two periods, where revenue in each period is an increasing function of the output of each agent and cost is the sum of the compensation payments to the agents. The agents are risk averse. Each agent's utility is additively separable across periods and within a period is additively separable between income and effort. We represent agents' limited access to the capital market by the extreme assumption that agents can neither save nor borrow (during the two periods), so they consume their payment in each period. The total utility of each agent is

$$U(I^1) - V(a^1) + \delta_A (U(I^2) - V(a^2)),$$

where I^t is the payment in period t and $\delta_A \in (0, 1]$ is the agents' discount factor. We assume

1 • $U : (0, \infty) \to (-\infty, \infty)$

2 • $U' > 0, U'' < 0, \lim_{I \to 0} U(I) = -\infty$

3 • $a^t \in [0, \infty), t = 1, 2$

4 • $V'(0) = 0, V'' > 0, \lim_{a \to \infty} V'(a) = \infty.$

Before period 1, the principal offers to each agent a compensation contract specifying the rules by which payments will be determined in both periods. We assume that if an agent accepts the contract, he is committed to stay with the firm for the two periods. We could equivalently assume that the skills that agents develop in the first period are so highly firm-specific that their outside opportunities in the second period will never induce them to quit. The two-period utility of each agent from not signing the contract is $\bar{U} > -\infty$. Note that under our assumptions, at the start of each period, the principal and the agents have common knowledge of the production functions and of one another's preferences. [1]

We assume that the agents' performance is monitored by an "observer", for whom the processing of information is costly and who consequently reports to the principal in each period only the rank order of agents' outputs. However, the principal can costlessly choose the criterion, called the "bias", that determines how the outputs are ranked: if the chosen bias in period t is c^t in i's favor, then i is declared the period-t winner if $x_i^t + c^t > x_j^t$ and j if $x_i^t + c^t < x_j^t$. We assume that the principal can commit himself, through the contract, to a level of first-period bias c^1, to a magnitude of second-period bias $|c^2|$, and to a rule determining, for each of the two first-period rankings, whether the second-period bias favors the first-period winner or loser. [2]

---

1. In this respect, the model is similar to ROGERSON'S [1985] and LAMBERT'S [1983] models of two-period incentive contracts for a single agent.
2. The result that it is optimal to use a strictly positive bias in the second contest in favor of the first-round winner would generalize if we relaxed the assumption that the magnitude of the second-period bias must be independent of which agent won the first contest. One rationale for this assumption is that the agents themselves may not be able to observe the rank order of their outputs (as in MALCOMSON [1984]). Then a scheme in which the magnitude of c^2 is independent of which agent the principal reports as the first-period winner is incentive compatible for the principal, whereas one in which $|c^2|$ varied according to the principal's report would not be: as Section 3 shows, equilibrium efforts are higher the smaller is $|c^2|$, so the principal would choose his report to bring about the lower value of $|c^2|$.

Bias in contests can be implemented in two different ways, depending upon why the observer reports only ordinal information about the agents' outputs to the principal. If the observer learns the actual values of outputs but can *transmit* ordinal information far more cheaply than cardinal information, then the principal can implement bias c in i's favor by instructing the observer to use asymmetric evaluation criteria, reporting j as the winner if and only if j's output exceeds i's by at least c. If ordinal information about outputs is far less costly for the observer to *gather* than cardinal information, the principal can control the bias by providing different inputs to the two agents' production functions or by differentiating the production functions themselves. The observer continues to observe and report only whether i's output is larger or smaller than j's, but these output values are perturbations on the x's above: the observer compares $x_i + v_i$ with $x_j + v_j$, where v_i and v_j are chosen by the principal so that $v_i - v_j$ equals the desired c. This second method of implementing bias can be used even when no individual in the organization receives information about outputs that is finer than rank-order information.

Asymmetric treatment of individuals, producing bias of the second type, can take numerous forms. Individuals may be assigned different tasks, placed in different environments, given different amounts of training or supervision, or supplied with different amounts of capital. For example, junior employees can be assigned to collaborate with senior colleagues who differ in their talent or motivation, and secretaries can be assigned equipment of different vintages. [3, 4]

We assume that the only way in which the contract can link second-period compensation to the first-period rank-order result is through the dependence of the assignment of second-period bias on the latter. An agent's second-period payment thus depends only on whether he wins or loses the (biased) second-period contest. It may well be easier for the principal to commit in advance to a magnitude of bias than to a plan

3. As an illustration of the contrast between the two different ways of implementing bias, consider first a promotion choice between two associate professors whose performances as assistant professors were different. The senior faculty can evaluate their publication records, but transmitting detailed evaluations to the dean may be difficult if the dean is not familiar with the relevant journals; it may be far easier to transmit simply a ranking of the two candidates and to incorporate bias by using asymmetric criteria in producing the ranking. Consider, in contrast, a department's decision about which of two graduating Ph.D. students to recommend more highly for jobs. Developing cardinal measures for the quality of theses can be very difficult even for the faculty, but a ranking is much easier to produce. The faculty can incorporate bias into the thesis evaluation process by basing its own input into a student's thesis on his previous performance in courses or exams. (This second method of implementing bias could also be used in the first example: the faculty could burden the associate professors with different levels of administrative responsibility and then simply report whose publication record was better.)
4. "Affirmative action" programs in the United States can be viewed as biasing the competition among workers through the use of asymmetric evaluation criteria; in contrast, "positive action" programs, such as those in Great Britain, increase the hiring or promotion prospects of some individuals by providing them with extra guidance or training (Commission for Racial Equality [1989]).

FIGURE 1

The Timing of Events.

- Principal offers two-period contract to agents, and agents accept.
- First-period bias c^1 implemented.
- Agents choose first-period efforts, a_i^1 and a_j^1.
- Interim evaluation: principal learns and reports biased ranking of first-period outputs, x_i^1 and x_j^1, and pays y^1+z^1 to winner and y^1-z^1 to loser.
- Second-period bias of magnitude $|c^2|$ implemented (sign of c^2 is contingent on first-period ranking).
- Agents choose second-period efforts, a_i^2 and a_j^2.
- Major promotion decision: principal learns and reports biased ranking of second-period outputs, x_i^2 and x_j^2, and pays y^2+z^2 to winner and y^2-z^2 to loser.

involving deferred payments. He can commit to $|c^2|$ through choices about organization design: for example, he can institutionalize different types of training or design (and give different titles to) jobs involving different packages of responsibilities.

A contract specifies the level of first-period bias, c^1; payments to the first-period winner and loser, y^1+z^1 and y^1-z^1, respectively; the magnitude of second-period bias, $|c^2|$, and the rule for assigning it, contingent on the first-period ranking; and payments to the second-period winner and loser, y^2+z^2 and y^2-z^2, respectively. (In each period, y^t is the average prize, and z^t measures the prize spread. Since $\lim_{I \to 0} U(I) = -\infty$, the principal will always choose $y^t > 0$ and $|z^t| < y^t$.) The interpretation is that the major promotion is awarded to the winner of the second contest; y^1+z^1 and y^1-z^1 represent salaries after the interim evaluation; and y^2+z^2 and y^2-z^2 represent salaries after the promotion decision (*see* Figure 1).

A Pareto optimal contact maximizes the discounted expected profit of the principal subject to the *ex ante* reservation utility constraints for the agents and the incentive compatibility constraints stemming from the agents' private choices of effort in each period. To express the incentive compatibility constraints, we assume that the principal can implement a set of effort levels $\{(a_i^1, a_j^1), (a_i^2, a_j^2)\}$ if and only if they form a subgame perfect equilibrium in the two-period game between the agents induced by the contract. Though our primary interest is the form that bias takes in Pareto optimal contracts, we prove our main result for the larger set of "cost-minimizing" contracts (*see* GROSSMAN and HART [1983]). A cost-minimizing contract for a given set of effort levels minimizes the discounted sum of compensation payments by the principal, subject to the reservation utility constraints and to the constraint that the effort levels constitute a subgame perfect equilibrium for the agents. Since in any Pareto optimal contract, the prizes and biases are cost-minimizing, given the efforts induced, our result on the use of bias in cost-minimizing contracts is necessarily valid for Pareto optimal ones as well.

3 The Second Period

Consider the Nash equilibrium behavior of the agents in the second period when the bias is c^2. (Our sign convention is that a positive bias favors i and a negative one favors j.) Given second-period prizes y^2+z^2 and y^2-z^2 and given a conjectured level of effort for j of a_j^2, i chooses a_i^2 to maximize

$$P(x_i^2+c^2>x_j^2)\,U(y^2+z^2)+(1-P(x_i^2+c^2>x_j^2))\,U(y^2-z^2)-V(a_i^2).$$

Given our assumptions, $\frac{\partial}{\partial a_i^2}P(x_i^2+c^2>x_j^2)$ does not depend on i's information about the first period. The first-order condition for an interior optimal a_i^2 is

$$\Delta U^2\frac{\partial}{\partial a_i^2}P(x_i^2+c^2>x_j^2)=V'(a_i^2),$$

where $\Delta U^2\equiv U(y^2+z^2)-U(y^2-z^2)>0$ as long as $z^2>0$. Now

$$\begin{aligned}\frac{\partial}{\partial a_i^2}P(x_i^2+c^2>x_j^2)&=\frac{\partial}{\partial a_i^2}P(\Delta\varepsilon^2>f^2(a_j^2,s^2)-f^2(a_i^2,s^2)-c^2)\\&=\frac{\partial}{\partial a_i^2}E_{s^2}[1-G^2(f^2(a_j^2,s^2)-f^2(a_i^2,s^2)-c^2)]\\&=E_{s^2}[g^2(f^2(a_j^2,s^2)-f^2(a_i^2,s^2)-c^2)f_1^2(a_i^2,s^2)],\end{aligned}$$

where E_{s^2} denotes the expectation with respect to s^2, and the second equality uses the independence of $\Delta\varepsilon^2$ and s^2. The first-order condition for an interior optimal a_j^2 is

$$\Delta U^2\frac{\partial}{\partial a_j^2}P(x_i^2+c^2<x_j^2)=V'(a_j^2),$$

and manipulations parallel to those above give

$$\frac{\partial}{\partial a_j^2}P(x_i^2+c^2<x_j^2)=E_{s^2}[g^2(f^2(a_j^2,s^2)-f^2(a_i^2,s^2)-c^2)f_1^2(a_j^2,s^2)].$$

Since $f_{11}^2\leqslant 0$ and $V''>0$, if the first-order conditions

$$\Delta U^2E_{s^2}[g^2(f^2(a_j^2,s^2)-f^2(a_i^2,s^2)-c^2)f_1^2(a_i^2,s^2)]=V'(a_i^2)$$
$$\Delta U^2E_{s^2}[g^2(f^2(a_j^2,s^2)-f^2(a_i^2,s^2)-c^2)f_1^2(a_j^2,s^2)]=V'(a_j^2)$$

have a solution, it is unique and has the form $a_i^2=a_j^2\equiv a^2$, so a^2 satisfies

(1) $$\Delta U^2g^2(-c^2)E_{s^2}[f_1^2(a^2,s^2)]=V'(a^2).$$

Since $V'(0)=0$, and $\lim_{a \to \infty} V'(a)=\infty$, there does exist an a^2 solving (1), and since $g^2(.)>0$ and $f_1^2(0, .)>0$, $a^2>0$.

The local second-order conditions for i and j, respectively, at $a_1^2=a_j^2=a^2$ are

$$\Delta U^2 E_{s^2}[-g^{2\prime}(-c^2)(f_1^2(a^2, s^2))^2+g^2(-c^2)f_{11}^2(a^2, s^2)]-V''(a^2)<0$$

$$\Delta U^2 E_{s^2}[g^{2\prime}(-c^2)(f_1^2(a^2, s^2))^2+g^2(-c^2)f_{11}^2(a^2, s^2)]-V''(a^2)<0$$

These are clearly satisfied if $c^2=0$, since $g^{2\prime}(0)=0$; and since $g^{2\prime}(.)$ is continuous, they will be satisfied at least for some interval of c^2 values with 0 in the interior. (The analysis below focuses on the effect of a small change in c^2 from 0.) It has been frequently noted in the tournament literature that local maxima of agents' expected utility functions might not be global maxima (*see* NALEBUFF and STIGLITZ [1983], O'KEEFE, VISCUSI, and ZECKHAUSER [1984]), and especially MOOKHERJEE [1989]). MOOKHERJEE'S arguments for unbiased tournaments apply to biased ones as well. Roughly, the global second-order conditions will be satisfied at $a_i^2=a_j^2=a^2$ if the distribution of $\Delta\varepsilon^2$ is sufficiently dispersed ($|g^{2\prime}(.)|$ sufficiently small) and/or the marginal disutility of effort rises sufficiently rapidly ($|V''(.)|$ sufficiently large). [5] In what follows, we assume that the global second-order conditions are satisfied at the unique solution to the first-order conditions.

Finally, there does not exist a Nash equilibrium in which an agent exerts no effort, since $V'(0)=0$, $g^2(.)>0$ and $f_1^2(0, .)>0$. Thus the unique Nash equilibrium in the biased second-period contest involves equal (strictly positive) effort levels for the two agents, with the common effort $a^2(c^2)$ solving (1).

Increasing the magnitude of the bias reduces the common effort level: since $g^2(.)$ is unimodal and symmetric about 0, (1) implies that a^2 is decreasing in $|c^2|$. For each agent, the marginal return to effort when effort levels are equal is proportional to the probability that $\Delta\varepsilon^2$ equals $-c^2$; this critical realization of $\Delta\varepsilon^2$ becomes less probable the further away from 0 it is. That effort levels are likely to fall as the degree of asymmetry between contestants rises has been noted before (*e.g.* LAZEAR and ROSEN [1981]). The important additional result for our analysis of bias in a two-period setting is the following: since $g^2(.)$ is symmetric about 0 and differentiable,

$$\left.\frac{da^2(c^2)}{dc^2}\right|_{c^2=0}=0.$$

Thus, a small increase in $|c^2|$ from 0 has no first-order effect on equilibrium efforts in the second period.

5. When $c^2=0$, the only conceivably profitable deviation is to a smaller, strictly positive effort; with bias in i's favor, j (but not i) could, in addition, conceivably benefit by an increase in effort. Both types of deviation are unprofitable under the conditions above.

4 The Optimal Two-Period Contract: Second-Period Bias Favors the First-Period Winner

To analyze the effect of bias c^1 in the first period, let ΔW^1 denote the difference (for each agent) between the overall expected utility from winning and losing the first contest (including the consequences from the use of second-period bias). An analysis parallel to that in Section 3 shows that the unique Nash equilibrium in the first period involves a common effort level a^1 which solves [6]

$$(2) \qquad \Delta W^1 g^1(-c^1) E_{s^1}[f_1^1(a^1, s^1)] = V'(a^1).$$

If the agents were risk neutral, a Pareto optimal contract would, in each period, set the bias equal to zero and specify the prize spread that induced the first-best levels of effort. Since the agents would not need to be compensated for the imposition of risk, the principal's expected payoff from this contract would be as high as if efforts could be directly observed and contracted upon.

4.1. Cost-Minimizing Contracts

We now analyze cost-minimizing contracts for the principal when the agents are risk averse. Specifically, we consider the principal's choice of first- and second-period prizes and first- and second-period biases to minimize his total discounted cost of implementing the common first-period effort level $a^1>0$ and the common second-period effort level $a^2>0$, subject to the reservation utility constraints of the agents. (Since $|c^2|$ is required to be independent of the first-period result and since, by (1), the equilibrium value of a^2 will be independent of the sign of c^2, a^2 will be independent of which agent wins the first contest.)

Our first main result is:

Proposition 1: In a cost-minimizing contract implementing $a^1>0$ and $a^2>0$, the first-period bias, c^1, is 0. Furthermore, if the magnitude of the second-period bias, $|c^2|$, is strictly positive, then that bias is, for both first-period rankings, assigned to favor the first-period winner.

Proof: *See* Appendix 1.

The intuition behind Proposition 1 is straightforward. First, reducing the magnitude of the first-period bias from a positive value to 0 raises first-period efforts by raising $g^1(-c^1)$ (*see* (2)) but has no effect on second-period efforts. Second, starting from a contract in which the second-period bias

6. The agents' second-order conditions will be satisfied under conditions analogous to those given in Section 3.

disadvantages the first-period winner, reducing its magnitude to 0 raises second-period efforts by raising $g^2(-c^2)$ (*see* (1)) and also raises first-period efforts by eliminating the future punishment for current success, thus increasing ΔW^1 (*see* (2)). Finally, the switch to a contract with $c^1=0$ and $|c^2|=0$ ensures that the agents' ex ante expected utilities are equal and therefore eliminates the rents paid to an agent who is advantaged ex ante.

4.2. The Effect of Introducing Bias in the Second Period

Given Proposition 1, the question that remains is whether, in a cost-minimizing contract, the principal uses no bias in the second period or commits to favor the first-period winner. We can now, without ambiguity, economize on notation by using c^2 in place of $|c^2|$, always taking $c^2 \geqslant 0$. We will analyze the cost-minimizing choice of first- and second-period prizes to implement $a^1>0$ and $a^2>0$, subject to the reservation utility constraints, given $c^1=0$ and $c^2 \geqslant 0$. We will show that, starting from $c^2=0$, a small increase in c^2 reduces the principal's total discounted cost.

The first-period winner wins the second contest with probability $p_w^2 \equiv 1-G^2(-c^2)$, since in equilibrium $a_i^2=a_j^2$; the first-period loser wins with probability $p_l^2 \equiv G^2(-c^2)$. Define

$$\Delta p^2(c^2) \equiv p_w^2 - p_l^2 = 1-2G^2(-c^2) \geqslant 0.$$

We can write ΔW^1, the difference between the overall rewards for winning and losing the first contest, as $\Delta U^1 + \delta_A \Delta p^2(c^2) \Delta U^2$ (where ΔU^1 is defined analogously to ΔU^2 and we again use the fact that $a_i^2=a_j^2$). Substitution into (2) (with $c^1=0$) gives

$$(3) \qquad [\Delta U^1 + \delta_A \Delta p^2(c^2) \Delta U^2]\, g^1(0)\, E_{s^1}[f_1^1(a^1, s^1)] = V'(a^1).$$

Equations (1) and (3) can be solved for ΔU^1 and ΔU^2 in terms of c^2, a^1, and a^2:

$$(4) \qquad \Delta U^1 = \frac{V'(a^1)}{g^1(0)\, E_{s^1}[f_1^1(a^1, s^1)]} - \frac{\delta_A \Delta p^2(c^2)\, V'(a^2)}{g^2(-c^2)\, E_{s^2}[f_1^2(a^2, s^2)]} \equiv R^1(c^2, a^1, a^2)$$

$$(5) \qquad \Delta U^2 = \frac{V'(a^2)}{g^2(-c^2)\, E_{s^2}[f_1^2(a^2, s^2)]} \equiv R^2(c^2, a^2)$$

$R^1(c^2, a^1, a^2)$ and $R^2(c^2, a^2)$ give the prize differences in utility terms necessary to implement efforts $a^1>0$ and $a^2>0$ for a given c^2. Clearly $R^2(c^2, a^2)>0$ and since $R^1(0, a^1, a^2)>0$, we can ensure that $R^1(c^2, a^1, a^2)>0$ by focusing on values of $c^2 \geqslant 0$ that are sufficiently small. Use the definitions of ΔU^1 and ΔU^2 to rewrite (4) and (5) as, respectively,

$$(6) \qquad U(y^1+z^1) - U(y^1-z^1) = R^1(c^2, a^1, a^2)$$

$$(7) \qquad U(y^2+z^2) - U(y^2-z^2) = R^2(c^2, a^2)$$

(6) implicitly defines a function $z^1=Z(y^1, R^1)$ for $y^1>0$. From the assumptions on $U(.)$, if $R^1>0$, then $0<Z(y^1, R^1)<y^1$ and $Z(.,.)$

is continuously differentiable and strictly increasing in both arguments. (7) implicitly defines z^2 by the same function: $z^2 = \mathrm{Z}(y^2, \mathrm{R}^2)$.

We now solve for the cost-minimizing contract that implements $a^1 > 0$ and $a^2 > 0$, given c^2. We solve

$$\min_{y^1, y^2} 2(y^1 + \delta_P y^2)$$

subject to

$$\frac{1}{2}\mathrm{U}(y^1 + \mathrm{Z}(y^1, \mathrm{R}^1(c^2, a^1, a^2))) + \frac{1}{2}\mathrm{U}(y^1 - \mathrm{Z}(y^1, \mathrm{R}^1(c^2, a^1, a^2))) - \mathrm{V}(a^1)$$
$$+ \delta_A\left[\frac{1}{2}\mathrm{U}(y^2 + \mathrm{Z}(y^2, \mathrm{R}^2(c^2, a^2)))\right.$$
$$\left. + \frac{1}{2}\mathrm{U}(y^2 - \mathrm{Z}(y^2, \mathrm{R}^2(c^2, a^2))) - \mathrm{V}(a^2)\right] \geqslant \bar{\mathrm{U}}.$$

The expression $2(y^1 + \delta_P y^2)$ is the principal's total discounted cost. The constraint is the common reservation utility constraint for the agents. Since $c^1 = 0$ and c^2 is assigned to favor the first-period winner, each agent has, before the first period, probability one-half of winning each contest, whatever the size of c^2. This class of contracts, therefore, despite being "unfair" *ex post* (in the second period), offers the agents identical expected utilities *ex ante*. The incentive compatibility constraints are incorporated through $\mathrm{Z}(y^1, \mathrm{R}^1(c^2, a^1, a^2))$ and $\mathrm{Z}(y^2, \mathrm{R}^2(c^2, a^2))$, which determine, for given positive average prize levels y^1 and y^2 and for given c^2, the monetary spreads necessary to induce efforts a^1 and a^2.

The principal's cost is clearly increasing in y^1 and y^2 and the same can be shown to be true of the agents' expected utility. Therefore, the reservation utility constraint will be binding at a solution. We write the Lagrangian (with multiplier $\lambda > 0$) as

$$\mathrm{L}(y^1, y^2, \lambda; c^2, a^1, a^2) = 2(y^1 + \delta_P y^2)$$
$$-\lambda\left\{\frac{1}{2}\mathrm{U}(y^1 + \mathrm{Z}(y^1, \mathrm{R}^1(c^2, a^1, a^2)))\right.$$
$$+ \frac{1}{2}\mathrm{U}(y^1 - \mathrm{Z}(y^1, \mathrm{R}^1(c^2, a^1, a^2))) - \mathrm{V}(a^1)$$
$$+ \delta_A\left[\frac{1}{2}\mathrm{U}(y^2 + \mathrm{Z}(y^2, \mathrm{R}^2(c^2, a^2)))\right.$$
$$\left.\left. + \frac{1}{2}\mathrm{U}(y^2 - \mathrm{Z}(y^2, \mathrm{R}^2(c^2, a^2))) - \mathrm{V}(a^2)\right] - \bar{\mathrm{U}}\right\}.$$

We now argue that, given (c^2, a^1, a^2), there exists a unique cost-minimizing contract, which we denote $(y^{1*}(c^2, a^1, a^2), y^{2*}(c^2, a^1, a^2))$, and that it is characterized by $\frac{\partial \mathrm{L}}{\partial y^1} = 0$ and $\frac{\partial \mathrm{L}}{\partial y^2} = 0$ (in conjunction with the binding constraint). First, in (y^1, y^2) space, the principal's iso-cost curves are straight lines with finite, strictly negative slope. Second, by the assumptions on U(.), the constraint can be shown to define a differentiable, strictly decreasing, and strictly convex function $y^2 = T(y^1)$. Furthermore, since $\bar{U} > -\infty$ and $\lim_{I \to 0} U(I) = -\infty$, the graph of T(.) lies entirely in the positive quadrant and has asymptotes of the form $y^1 = k^1 \geqslant 0$ and $y^2 = k^2 \geqslant 0$. It follows that the unique point of tangency between an isocost line and the graph of T(.) is the unique cost-minimizing contract, $(y^{1*}(c^2, a^1, a^2), y^{2*}(c^2, a^1, a^2))$, and that $y^{1*} > 0$ and $y^{2*} > 0$.

Define $C^*(c^2, a^1, a^2)$ as the minimized cost to the principal of implementing efforts a^1 and a^2, given that c^2 is assigned to favor the first-period winner. Since the principal's cost and the agents' expected utility are both continuously differentiable functions of y^1, y^2 and c^2, it follows by using the implicit function theorem on the first-order conditions that y^{1*} and y^{2*}, and hence C^*, are differentiable in c^2.

To calculate $\frac{\partial C^*}{\partial c^2}$, use the envelope theorem, and since c^2 enters the Lagrangian directly only through

$$Z(y^1, R^1(c^2, a^1, a^2)) \quad \text{and} \quad Z(y^2, R^2(c^2, a^2))$$

in the constraint,

$$(8) \quad \frac{\partial C^*}{\partial c^2}(c^2, a^1, a^2) = -\frac{1}{2}\lambda \Big\{ [U'(y^{1*} + Z(y^{1*}, R^1(c^2, a^1, a^2)))$$
$$- U'(y^{1*} - Z(y^{1*}, R^1(c^2, a^1, a^2)))] \frac{\partial Z}{\partial R}(y^{1*}, R^1) \frac{\partial R^1}{\partial c^2}(c^2, a^1, a^2)$$
$$+ \delta_A [U'(y^{2*} + Z(y^{2*}, R^2(c^2, a^2)))$$
$$- U'(y^{2*} - Z(y^{2*}, R^2(c^2, a^2)))] \frac{\partial Z}{\partial R}(y^{2*}, R^2) \frac{\partial R^2}{\partial c^2}(c^2, a^2) \Big\}$$
$$= -\frac{1}{2}\lambda \Big\{ D^1 \frac{\partial Z}{\partial R}(y^{1*}, R^1) \frac{\partial R^1}{\partial c^2}(c^2, a^1, a^2)$$
$$+ D^2 \frac{\partial Z}{\partial R}(y^{2*}, R^2) \frac{\partial R^2}{\partial c^2}(c^2, a^2) \Big\},$$

where $D^t \equiv U'(y^{t*} + Z(y^{t*}, R^t)) - U'(y^{t*} - Z(y^{t*}, R^t))$ for $t = 1, 2$. Differentiating (4) and (5) with respect to c^2 and using $a^2 > 0$ yields

$$(9) \quad \frac{\partial R^1}{\partial c^2}(c^2, a^1, a^2) = -\frac{\delta_A V'(a^2)}{E_{s^2}[f_1^2(a^2, s^2)]} \times \left[\frac{2(g^2(-c^2))^2 + \Delta p^2(c^2) g^{2\prime}(-c^2)}{(g^2(-c^2))^2} \right] < 0, \qquad \forall c^2 \geqslant 0$$

$$(10) \quad \frac{\partial R^2}{\partial c^2}(c^2, a^2) = \frac{V'(a^2)}{E_{s^2}[f_1^2(a^2, s^2)]} \left[\frac{g^{2\prime}(-c^2)}{(g^2(-c^2))^2} \right] \geqslant 0, \qquad \forall c^2 \geqslant 0$$

Evaluating these derivatives at $c^2=0$ gives

$$\frac{\partial R^1}{\partial c^2}(0, a^1, a^2)<0 \qquad \text{and} \qquad \frac{\partial R^2}{\partial c^2}(0, a^2)=0, \quad \text{since } g^{2\prime}(0)=0.$$

Since $R^1(0, a^1, a^2)>0$ and $R^2(0, a^2)>0$, $D^t<0$ for $t=1, 2$ when evaluated at $c^2=0$. From (6) and (7), $\frac{\partial Z}{\partial R}(y^{t*}, R^t)>0$ for $t=1, 2$. Therefore, (8) shows that for all $a^1>0$ and $a^2>0$,

$$\frac{\partial C^*}{\partial c^2}(0, a^1, a^2)<0.$$

Hence, given Proposition 1, we have shown

PROPOSITION 2: A cost-minimizing contract implementing any $a^1>0$ and $a^2>0$ commits the principal to favor the first-period winner by a strictly positive amount in the second period. Therefore, as long as the optimal efforts in a Pareto optimal contract are strictly positive, a Pareto optimal contract uses strictly positive second-period bias in favor of the first-period winner.

For the intuition behind Proposition 2, consider how an increase in c^2 affects the agents' expected utility, when y^1 and y^2 are held fixed at their optimal levels and z^1 and z^2 are adjusted so the agents continue to choose a^1 and a^2. An increase in c^2 lowers the marginal effect of effort on the probability of winning the second contest, and keeping a^2 fixed requires offsetting this change by an increase in ΔU^2 (*see* (10) and (1)). On the other hand, an increase in c^2 leads to a reduction in ΔU^1, because it increases both ΔU^2 and $\Delta p^2(c^2)$, and hence increases the component of ΔW^1 representing the second-period reward to winning the first contest; the first-period reward must therefore be reduced to keep a^1 unchanged (*see* (9) and (3)). With the average prizes y^1 and y^2 fixed, these changes in ΔU^1 and ΔU^2 are accomplished by changes in the prize spreads z^1 and z^2; the increase in ΔU^2 therefore lowers the risk averse agents' expected utility, while the reduction in ΔU^1 raises it. Starting from $c^2=0$, however, the increase in ΔU^2 is of only second order, since the reduction in effort a^2 for a given ΔU^2 is of only second order (as shown in Section 3). But the reduction in ΔU^1 is of first order, since the increase in $\Delta p^2(c^2)$, and therefore in the effort a^1 for a given ΔU^1, is of first order. Therefore, the only first-order effect of an increase in c^2 from 0 is an increase in the agents' expected utility from the first period. Since the principal will optimally hold the agents to their reservation utility, introducing a small second-period bias in favor of the first-period winner produces a first-order reduction in the principal's expected cost. [7, 8]

7. This result remains valid even if the incentives for second-period effort are provided partly by the adjustable prize spread ΔU^2 and partly by an exogenously given reward, representing future opportunities available only to the promoted agent.

8. As either a^1 or a^2 approaches 0, $\frac{\partial C^*}{\partial c^2}(0, a^1, a^2)$ approaches 0. Inducing $a^1=0$ with $c^2=0$ requires no imposition of risk in the first period, so raising c^2 generates no risk-reduction benefit. Inducing $a^2=0$ requires no prize spread in the second period, so raising c^2 from 0 generates no increase in first-period incentives and therefore no risk-reduction benefit.

Hence it is optimal for the principal to commit to spreading rewards over time, by rewarding the first-period winner not only with a higher salary but also with a job in which his future promotion probability is higher. A correlation between early and later success is therefore induced, even though first-period success is due, in equilibrium, entirely to transitory random shocks.

Since bias in the second period reduces effort incentives at that point, it will be implemented by the principal only if he is committed to doing do, for example, through the job structures he has designed. The organizational sociologists BARON and BIELBY [1986] have documented the significant extent to which organizations "fragment" work through the proliferation of job titles, "making finer distinctions among work roles than are required simply on the basis of job content". These differences in job titles may be the organization's way of formally identifying the use of bias. This hypothesis is consistent with BARON and BIELBY'S finding that the proliferation of job titles is more extreme in firms in which workers' skills are more highly firm-specific: such firms should find it easier to introduce bias, and hence reap the incentive benefits, without driving away those whom the bias disadvantages.

5 Discussion

We now discuss two simple settings in which the agents are not identical and extend our result on the value of committing to conduct an unfair contest in the second period.

5.1. Unobservable Differences in Promotion-Relevant Skills in the Second Period

Consider a setting in which the principal uses the contests both to provide incentives and to learn about differences in the agents' skills. In the first period, the agents perform a very routine task in which their outputs are insensitive to ability, depending only on effort, as in the original model. However, the agents also automatically develop firm-specific skills. The levels of skills developed by i and j, η_i^2 and η_j^2, are *ex ante* uncertain and independent of first-period efforts, because they reflect innate, unobservable ability. The common prior beliefs of the principal and agents are symmetric with respect to η_i^2 and η_j^2 (the agents have equal potential) and assign $\Delta\eta^2 \equiv \eta_i^2 - \eta_j^2$ a unimodal density. Post-promotion profits are maximized by promoting the more skilled agent, and the only information about skills that becomes available to either the principal or the agents before the promotion decision is the (biased) ranking of second-period outputs. The difference in second-period outputs, $\Delta x^2 \equiv x_i^2 - x_j^2$, depends on $\Delta\eta^2$ according to

(11) $$\Delta x^2 = f^2(a_i^2, s^2) - f^2(a_j^2, s^2) + \Delta\varepsilon^2 + \Delta\eta^2,$$

where $\Delta\varepsilon^2$ and $\Delta\eta^2$ are independent. This formulation of the technology preserves the result that the agents' second-period equilibrium efforts are equal, no matter what bias is used. Furthermore, (11) is a natural extension, to situations with moral hazard, of the framework used in MEYER [1991] to examine the value of biased contests in learning about skills.

In this setting, we assume that the principal chooses a contract from class analyzed above to maximize the discounted sum of expected pre-promotion and expected post-promotion profit, subject to implementing given efforts $a^1>0$ and $a^2>0$. The second-period bias affects the former component of profit through the costs of providing incentives and the latter by influencing the information about relative skills provided by the second contest.

Consider first expected pre-promotion profit. Let $h^2(.)$ denote the density function of $\Delta\varepsilon^2+\Delta\eta^2$. Given a bias of c^2, the unique Nash equilibrium in the second contest involves a common effort level a^2 which solves

$$\Delta U^2 h^2(-c^2)\,E_{s^2}[f_1^2(a^2, s^2)] = V'(a^2).$$

Given our assumptions on $\Delta\eta^2$ and $\Delta\varepsilon^2$, $h^2(.)$ is symmetric about 0, unimodal, and has a stationary point at 0, so a small increase in $|c^2|$ has no first-order effect on a^2 for a given ΔU^2. The function h^2 plays exactly the same role here as g^2 played in Sections 3 and 4: here the "effective noise term" is $\Delta\varepsilon^2+\Delta\eta^2$, instead of just $\Delta\varepsilon^2$. Our previous analysis therefore implies that the expected pre-promotion profit from implementing efforts $a^1>0$ and $a^2>0$ is maximized by setting $c^1=0$ and committing to favor the first-period winner by $|c^2|>0$.

Now consider the effect of c^2 on expected post-promotion profit. No matter what the common level of a^2, (11) implies that $\Delta x^2=\Delta\varepsilon^2+\Delta\eta^2$. Since, at the start of the second period, the principal's beliefs are symmetric with respect to η_i^2 and η_j^2, the optimal promotion rule is, for any c^2, to promote the second-period winner. With this rule, since the distribution of $\Delta\varepsilon^2$ is symmetric about 0 and continuous, expected post-promotion profit is symmetric about $c^2=0$ and (as shown in Appendix 2) has a stationary point there. Hence a small increase in $|c^2|$ from 0 to favor the first-period winner has only a second-order effect on expected post-promotion profit. With *ex ante* indistinguishable agents, a symmetric production environment, and continuously distributed noise, the payoff from the promotion decision is, to first order, unaffected by the introduction of bias.

Therefore, in a contract that maximizes the discounted sum of expected pre-promotion and expected post-promotion profit, subject to implementing given efforts $a^1>0$ and $a^2>0$, the principal commits to use strictly positive second-period bias in favor of the first-period winner. Bias in the second contest can reduce both second-period incentives and the value of the information provided about relative skills but, starting from zero bias, each of these effects is of second order, while the increase in first-period incentives is of first order. The principal's concern with learning in this setting does not alter the conclusion of the earlier analysis: it is optimal to structure the second contest so the first-period winner is strictly more likely than his rival to win and be promoted.

5.2. Observable Differences in Productivity in Pre-Promotion Tasks

Consider a different scenario, in which there are, *ex ante, commonly known* productive differences between the agents in the entry-level jobs. Specifically, let Δx^2 be given by (11) and $\Delta x^1 \equiv x_i^1 - x_j^1$ by

$$\Delta x^1 = f^1(a_i^1, s^1) - f^1(a_j^1, s^1) + \Delta\varepsilon^1 + \Delta\eta^1,$$

where $\Delta\eta^1$ and $\Delta\eta^2$ are commonly known constants. Suppose that after two periods, both agents will be reassigned, though only one will be promoted, and that $\Delta\eta^1$ and $\Delta\eta^2$ are uninformative about how they will perform after reassignment. Thus $\Delta\eta^1$ and $\Delta\eta^2$ may represent differences in skills that the agents use only in the entry-level positions or naturally arising differences in the tasks performed in these positions.

In this setting, as in the original model, bias in the first two periods has no effect on profits after the promotion decision, so the use of bias will be determined entirely by the costs of providing incentives. Recall the sign convention that a positive bias favors i and a negative one favors j. The unique Nash equilibrium efforts a^1 in the first contest, given a bias of c^1, solve

$$\Delta W^1 g^1(-c^1 - \Delta\eta^1) E_{s^1}[f_1^1(a^1, s^1)] = V'(a^1),$$

where ΔW^1 depends upon the rule determining the level of bias in the second contest. The unique Nash equilibrium efforts a^2 in the second contest, given a bias of c^2, solve

$$\Delta U^2 g^2(-c^2 - \Delta\eta^2) E_{s^2}[f_1^2(a^2, s^2)] = V'(a^2).$$

Consider the class of contracts specifying first-period bias $c^1 = -\Delta\eta^1 + d^1$ and second-period bias $c_i^2 = -\Delta\eta^2 + d^2$ if i wins and $c_j^2 = -\Delta\eta^2 - d^2$ if j wins. It is easy to adapt the proofs of Propositions 1 and 2 to show that the principal minimizes the cost of implementing $a^1 > 0$ and $a^2 > 0$ by choosing $d^1 = 0$ and $d^2 > 0$. Thus the first-period bias should exactly offset the productive difference $\Delta\eta^1$ and so make the contest fair, but the second-period bias should deviate from $-\Delta\eta^2$ by rewarding the first-period winner with a strictly higher probability of winning again than in a fair contest.

In general, when the performance of initially indistinguishable employees in a sequence of contests is informative about both their efforts and their abilities, then in choosing how to use bias, the organization may need to trade off the incentive and learning benefits. When the first-period result is informative about relative abilities, the effort-maximizing bias in the second period will typically favor the first-period loser, offsetting the probable difference in abilities. Sections 4 and 5 suggest that the principal's pre-promotion profit will be increased by committing to deviate in the second period from the effort-maximizing bias to supplement the reward to the first-period winner. Still, the second-period bias that maximizes pre-promotion profits could, relative to no bias at all, disadvantage the first-period winner. On the other hand, the analysis in MEYER [1991] implies that post-promotion profits will be maximized by setting the second-period bias to favor the first-period winner relative to no bias, *i.e.* to reinforce the probable difference in abilities. Hence the choice of second-period bias may have opposite implications for employee incentives and for learning by the organization, and which employee should be favored may depend on the relative sensitivity to bias of pre-promotion and post-promotion profits. Alternatively, the use of bias may in practice depend on the incentives of the individual who controls the conditions under which employees compete. If bias is controlled by the supervisor of junior workers, who is compensated according to their productivity before promotion, then it is likely to be dictated by incentive considerations. If, instead, bias is controlled by the future manager of the promoted worker, then learning considerations are more likely to dominate.

APPENDIX 1

Proof of Proposition 1

Step (i): Start with an arbitrary contract, A, which implements $a^1>0$ and $a^2>0$ and satisfies both agents' reservation utility constraints but in which one or both of the following are true: (I) $c^1 \neq 0$; (II) $|c^2|>0$ and for at least one first-period ranking, the first-period loser is favored by $|c^2|$. Replace contract A by B, which specifies the same first- and second-period prizes but sets $c^1=0$ and $|c^2|=0$.

Step (ii): Under contract B, each agent has, in equilibrium, *ex ante* probability one-half of winning each contest, so the expected value of the utility from monetary prizes is equal for the two agents. If contract A makes this expected value lower for one agent than the other, then the switch from A to B raises this expected value for the initially worse-off agent – since the sum over the agents of these expected values is the same under A and B.

Step (iii): By (1), second-period efforts are independent of the first-period ranking and are (weakly) larger under contract B than under A. Replace contract B by C, which differs from B only in that z^2 is (weakly) reduced so that the product $\Delta U^2 g^2(-c^2)$ is equal under A and C, and therefore a^2 is equal under A and C. Since by assumption $a^2>0$ under A, the value of z^2 under C must be strictly positive.

Step (iv): Under contract A, $\Delta W^1 \leqslant \Delta U^1 \equiv U(y^1+z^1)-U(y^1-z^1)$ (*see* condition (II) in Step (i)), whereas under C, $\Delta W^1 = \Delta U^1$. Also, $g^1(-c^1)$ is (weakly) larger under C than under A (*see* condition (I)). Therefore, by (2), first-period efforts are (weakly) larger under C than under A. Replace C by D, which differs from C only in that z^1 is (weakly) reduced so that the product $\Delta W^1 g^1(-c^1)$ is equal under A and D, and therefore a^1 is equal under A and D. Since by assumption $a^1>0$ under A, the value of z^1 under D must be strictly positive.

Step (v): Since, by hypothesis, contract A satisfies one or both of (I) and (II) in Step (i), either one or both of z^1 and z^2 are *strictly* smaller under D than under A.

Step (vi): Overall, the switch from contract A to D leaves a^1 and a^2 unchanged. The principal's total discounted cost is also unaffected, since total payments in each period ($2y^1$ and $2y^2$) are unchanged. If one agent's expected utility under contract A is lower than the other's, the switch from A to D increases the expected utility of the initially worse-off agent, in three ways: (*a*) Step (ii) showed that, for given, y^1, z^1, y^2, z^2, the change in the success probabilities in the two contests increases his expected value of the utility from monetary prizes; (*b*) given probability one-half under D of winning the second contest, the reduction in z^2 (Step (iii)) with y^2 fixed raises expected utility given risk aversion; and (*c*) given probability one-half under D of winning the first contest, the reduction in z^1 (Step (iv)) with y^1 fixed raises expected utility given risk aversion. By Step (v), one or both of the increases in (*b*) and (*c*) are strictly positive, and these increases are present (for both agents) even if, under A, the agents' expected utilities are equal. Since D gives them equal expected utilities, the switch from A to D slackens the tighter of the two reservation utility constraints.

Step (*vii*): Therefore, starting from contract D, the principal can reduce his total discounted cost, while continuing to implement a^1 and a^2 and to satisfy the reservation utility constraints, by slightly reducing y^1 and y^2 and simultaneously adjusting z^1 and z^2 to keep ΔU^1 and ΔU^2 unchanged. This final modification to D yields a contract that strictly dominates A. □

APPENDIX 2

For the model of Section 5.1, we prove that when the principal uses the optimal rule, "promote the second-period winner", expected post-promotion profit has a stationary point at $c^2=0$. Let $l(\eta_i^2, \eta_j^2)$ be the symmetric joint density function of η_i^2 and η_j^2. Given skills η_i^2 and η_j^2, let the principal's profit if i is promoted be $\pi^i(\eta_i^2, \eta_j^2)$ and if j is promoted be $\pi^j(\eta_i^2, \eta_j^2)$. The principal's profit should depend only on the agents' skills and not on their identities, so $\pi^i(a, b)=\pi^j(b, a)$ for all (a, b). Since the agents' second-period efforts are equal, (11) implies that i is promoted if $\Delta\varepsilon^2+\Delta\eta^2+c^2\geqslant 0$ and j is promoted if $\Delta\varepsilon^2+\Delta\eta^2+c^2<0$. Hence expected post-promotion profit, as a function of c^2, is

$$\Pi(c^2)\equiv\int\int[\pi^i(a, b)\,\mathrm{P}(\Delta\varepsilon^2\geqslant -a+b-c^2) + \pi^j(a, b)\,\mathrm{P}(\Delta\varepsilon^2< -a+b-c^2)]\,l(a, b)\,da\,db,$$

where a and b are the variables of integration for η_i^2 and η_j^2, respectively.

$$\frac{\partial\Pi}{\partial c^2}(0)=\int\int[\pi^i(a, b)-\pi^j(a, b)]\,g^2(-a+b)\,l(a, b)\,da\,db$$

Using the symmetry of $l(a, b)$, $\frac{\partial\Pi}{\partial c^2}(0)$ can be written as

$$\int_{a\geqslant b}\int_b\{[\pi^i(a, b)-\pi^j(a, b)]\,g^2(-a+b)+[\pi^i(b, a) - \pi^j(b, a)]\,g^2(-b+a)\}\,l(a, b)\,da\,db.$$

But

$$[\pi^i(a, b)-\pi^j(a, b)]\,g^2(-a+b)+[\pi^i(b, a)-\pi^j(b, a)]\,g^2(-b+a) = g^2(-a+b)\{[\pi^i(a, b)-\pi^j(b, a)]+[\pi^i(b, a)-\pi^j(a, b)]\}=0,$$

since $g^2(.)$ is symmetric about 0 and $\pi^i(a, b)=\pi^j(b, a)$ for all (a, b). Hence $\frac{\partial\Pi}{\partial c^2}(0)=0$. □

• References

BARON, J. and BIELBY, W. (1986). – "The Proliferation of Job Titles in Organizations", *Administrative Science Quarterly*, Vol. 31, pp. 561-586.

COMMISSION FOR RACIAL EQUALITY (1989). – "Positive Action and Equal Opportunity in Employment", London.

GILSON, R. J. and MNOOKIN, R. H. (1989). – "Coming of Age in a Corporate Law Firm: The Economics of Associate Career Patterns", *Stanford Law Review*, Vol. 41, pp. 567-595.

GROSSMAN, S. and HART, O. (1983). – "An Analysis of the Principal-Agent Problem", *Econometrica*, Vol. 51, pp. 7-46.

KAHN, C. and HUBERMAN, G. (1988). – "Two-sided Uncertainty and 'Up-or-Out' Contracts", *Journal of Labor Economics*, Vol. 6, pp. 423-444.

KANTER, R. M. (1977). – *Men and Women of the Corporation*. New York: Basic Books.

LAFFONT, J.-J. and TIROLE, J. (1988). – "Repeated Auctions of Incentive Contracts, Investment, and Bidding Parity with an Application to Takeovers", *Rand Journal of Economics*, Vol. 19, pp. 516-537.

LAMBERT, R. A. (1983). – "Long-Term Contracts and Moral Hazard", *Bell Journal of Economics*, Vol. 14, pp. 441-452.

LAZEAR, E. P. and ROSEN, S. (1981). – "Rank-Order Tournaments as Optimum Labor Contracts", *Journal of Political Economy*, Vol. 89, pp. 841-864.

MALCOMSON, J. (1984). – "Work Incentives, Hierarchy, and Internal Labor Markets", *Journal of Political Economy*, Vol. 92, pp. 486-507.

MEYER, M. A. (1991). – "Learning from Coarse Information: Biased Contests and Career Profiles", *Review of Economic Studies*, Vol. 58, pp. 15-41.

MILGROM, P. and ROBERTS, J. (1988). – "An Economic Approach to Influence Activities in Organizations", *American Journal of Sociology*, Vol. 94, Supplement, pp. S154-S179.

MOOKHERJEE, D. (1989). – "Rank Order Competition and Incentives: An Organizational Perspective", *Mimeo*, Stanford Graduate School of Business, February.

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

O'KEEFE, M., VISCUSI, K. and ZECKHAUSER, R. (1984). – "Economic Contests: Comparative Reward Schemes", *Journal of Labor Economics*, Vol. 2, pp. 27-56.

PRENDERGAST, C. (1989). – "Career Development and the Signaling Effect of Wages and Promotion", *Mimeo*, Nuffield College, Oxford, October.

ROGERSON, W. (1985). – "Repeated Moral Hazard", *Econometrica*, Vol. 53, pp. 69-76.

ROSENBAUM, J. (1984). – *Career Mobility in a Corporate Hierarchy*. Academic Press.

WALDMAN, M. (1990). – "Up-or-Out Contracts: A Signaling Perspective", *Journal of Labor Economics*, Vol. 8, pp. 230-250.