Title
Tournaments, Termination Schemes, and Forcing Contracts

Author
Bhattacharya, Sudipto

Publication Date
1985-06-01

Peer reviewed

TOURNAMENTS, TERMINATION SCHEMES, AND FORCING CONTRACTS

Sudipto Bhattacharya

June 1985

Research Papers in Economics No. 85-7

Acknowledgments: This paper is a revised version of a segment of Bhattacharya [1983], originating in a London School of Economics mimeograph of Autumn 1982. The Business Schools at Stanford, Harvard, and Berkeley (Program in Finance), together with the Batterymarch Financial Management fellowship, provided interim financial and secretarial support. Numerous colleagues have been generous with their help. I would particularly like to thank John Geanakoplos, Dilip Mookherjee, Barry Nalebuff, Jose Scheinkman, Joseph Stiglitz, and an anonymous associate editor for stimulating comments and suggestions, while retaining responsibility for any errors or ambiguities.

ABSTRACT

The performance of tournament incentive schemes is compared with that of schemes based on termination threats, in an environment with two-sided moral hazard for both principal and agent, and "moving support" monitoring. The disadvantages of the former scheme that arise from having to eliminate "collusive" multiple equilibria among agents are contrasted with the "involuntary unemployment" deadweight-loss of termination-based schemes. Tournament-type payoff structures are shown to be inessential in obtaining asymptotic first-best optimal resolutions, given moral hazard on the part of agents alone, and monitoring noise distributions with appropriate lower-tail properties.

1. Introduction

In recent years, following upon the early work on moral hazard by Spence and Zeckhauser [1971] and Mirrlees [1979], there has been much research activity devoted to the generalization and extension of the basic results regarding optimal reward schemes, to the scenarios of multiagent incentive contracts and incentive effects of termination schemes. A significant segment of the former literature has focused on the properties of rank-order (Tournament) incentive schemes, in which agents' rewards are determined simply by the ordinal rank of their observable output levels in relation to those of other agents, e.g., Lazear and Rosen [1981], Bhattacharya [1982], Nalebuff and Stiglitz [1983], and Green and Stokey [1983]. More general properties of multiagent incentive schemes have been discussed by Holmström [1982], and Mookherjee [1984] in particular. Incentive schemes (which motivate agents to provide unobservable effort) that are based on the threat of termination of contractual relationships, rather than contemporaneous adjustment of agents' payoffs contingent on observed outcomes, have been developed in Shapiro and Stiglitz [1984].

In much of the work on multiagent incentive schemes, e.g., Bhattacharya [1982], the rationale for linking one agent's reward to another agent's performance has been that of adjusting for the effects of correlated productivity shocks that impinge on the performance of both.[1] Our focus in most part in this paper is, in contrast, on a rationale for linkages across agents in incentive payments that is not dependent on (appropriately generalized notions of) correlation in observations of their productivities, conditional on effort levels. This rationale pertains to a scenario in which there is moral hazard on the part of both principal and agent, in that while the latter observe their effort choices asymmetrically, the former in turn observe the resulting performance indicators (monitors) asymmetrically ex post, or at least it is the case that independent third-party verification of such monitored information is infeasible. Carmichael [1983] was the first to note, in the context of seniority-based incentive schemes, that particular forms of multiagent linkages imply that the principal's total wage bill, aggregated across agents, would be independent of his (privately known) reported monitoring information. Thus, it might be (weakly) incentive-compatible for the principal to adhere to prespecified linkages between agents' performance levels and resulting payoffs.[2]

Independently of Carmichael's work Shapiro and Stiglitz [1984] have noted that, in a scenario with two-sided moral hazard as described above, reward schemes based on the threat of termination contingent on "bad" observations of performance have a similar incentive-compatibility property for the principal. In competitive labor markets, the equilibrium wage level equals the marginal product of labor, and thus principals are ex post indifferent (in terms of their pecuniary payoffs) about firing a worker. Hence, the ex ante incentive effects of firing threats given bad performance prevail in implementing a wage-employment scheme. There is, however, a (social) cost to operating such incentive schemes, particularly when individuals in the pool of unemployed workers are anonymous, i.e., no individual-specific ability differences are signaled by the fact of having been fired. At any point in time, there needs to be a pool of involuntarily unemployed workers, for whom the going wage strictly exceeds the marginal utility of the leisure obtained from non-employment, so that there is a (private) cost to being terminated from an employment relationship.[3] In an ex ante sense, one may also view this cost in (private) terms of a randomization in being (gainfully) employed, at a wage strictly exceeding the utility of leisure.

At first glance, it would appear that the rationale for multiagent (nontermination) incentive schemes offered by Carmichael and ourselves, would largely eliminate the desirability of termination schemes, since the former type of scheme does not require a pool of (involuntarily) unemployed workers. There are, however, two qualifications to such an intuition. First, it is easy to show, along the lines of Holström [1982], that in the absence of correlations in the random errors in measuring workers' productivities, multiagent schemes impose more welfare-decreasing noise on risk-averse workers, for a given level of induced effort. Termination schemes, in contrast, are based on agent-specific performance measurements, and thus are not subject to this problem.[4]

However, as Green and Stokey [1983] have noted for a simplified environment, this difficulty with multiagent schemes may not be relevant if the number of agents is sufficiently large, and the monitoring structure allows for an equally large set of partitions to accommodate the resulting cardinality of performance rank-orders. While the Green-Stokey result, on the second-best optimality of tournaments with many agents, may not be entirely relevant to the small numbers (and inherently coarse monitoring) scenario of hierarchical incentive schemes that the Tournaments literature first sought to address (see also Bhattacharya and Guasch [1985]), it is pertinent to the specific scenario we shall deal with below.

A second difficulty with multiagent schemes, discussed in the work of Mookhergee [1984], has to do with the existence of multiple Nash equilibria in the subgame among agents that is generated by the (Stackelberg leader) principal's incentive scheme. Specifically, given a reward scheme, the equilibrium preferred by the principal may be Pareto-dominated for agents by another Nash equilibrium involving a lower set of effort levels. Thus, to be relevant to empirical reality and prediction, the theory of optimal design of incentive schemes must account for the means by which principals (employers) would design reward structures to eliminate such (potentially) "collusive" Nash equilibria among agents, i.e., the notion of Nash implementation must incorporate the nonexistence of other Pareto-dominating Nash equilibria across agents at the very least. As we show below, this consideration is of great importance even for the simplest multiagent schemes like rank-order tournaments, which also resolve the principal's moral hazard problem. This is in contrast to the conjecture of Holmström [1982], who asserts that tournament schemes generate "zero-sum games" across agents, and hence cannot have Pareto-ordered Nash equilibria, an assertion which amounts to a confusion with their constant-sum (across agents) pecuniary payoff (gross of disutility of effort) property.

This (potential) collusion aspect to multiagent schemes, detailed above, is exceedingly relevant to the comparison of their induced welfare attainment properties, vis-à-vis that of termination-based schemes. As shown below, this is the case in the presence of two-sided moral hazard, even if the number of agents is "large," i.e., countably infinite. Indeed, potential distortions caused by this problem persist even when the monitoring structure has a moving support property, i.e., the support of the distribution of an agent's performance monitor is altered by the effort level chosen, which is a case in which single-agent "forcing contract" incentive schemes (in the absence of principal's moral hazard) may attain the first-best allocation. The essence of this difficulty with multiagent schemes arises from the following interaction across the problems of (i) avoiding collusive Nash equilibria that are worse (better) for the principal (agents), and (ii) principal's moral hazard in implementing incentive schemes, as detailed below.

It is shown in section 2 below, in the context of a moving support monitoring structure that is in the spirit of the Shapiro-Stiglitz [1984] model, that the principal needs the power to penalize a sufficiently large fraction of workers in order to eliminate the "bad" Nash equilibrium of shirking by all workers. However, given this power, the principal will use it even when the "good" Nash equilibrium is obtained, and penalize (pay low wages to) workers in a purely randomized fashion. It is shown in section 2 that, nevertheless, for risk-neutral workers[5] it is feasible to design a tournament scheme that strictly dominates the termination-based scheme of Shapiro-Stiglitz in the resulting welfare attainment. However, for sufficiently risk-averse workers, the cost (welfare loss) of such ex post randomization may well exceed that of ex ante (randomized) unemployment in the Shapiro-Stiglitz scheme; an example of this phenomenon is provided.

The above-mentioned set of tradeoffs, between tournament and termination schemes, is crucially dependent on an additional assumption, namely that agents' (von Neumann-Morgenstern) utility functions are bounded below, so that an infinitesimally small probability of (feasibly) low wages is not sufficient to induce the desired effort level. (Similarly, for moving-support individualistic schemes, the probability of detecting shirking needs to be high enough.) A classic theorem in the moral hazard literature, due to Mirrlees [1979], shows that for agents with utility functions that are unbounded below, the first-best effort and expected utility levels are attainable arbitrarily closely even without moving-support monitoring, given some (plausible) conditions on the probability-distribution of noisily monitored output/effort levels.

Nalebuff and Stiglitz [1983] have recently proved an analogous result on asymptotic first-best attainment, given a large number of agents, partially correlated productivity/monitoring noises, and a different set of distributional conditions on the probability of coming in last in a (many-person) tournament scheme. In section 3 and the Appendix, we show that the Nalebuff-Stiglitz conditions imply those in Mirrlees [1979], and that the Mirrlees incentive scheme is extendable to the correlated-shock case.[6] Thus, in some sense, the "primacy" of the two-sided moral hazard rationale for tournament schemes is further established. In section 4, which concludes the paper, we discuss suggestions for further research along the lines developed in this paper, as well as on other issues associated with tournament schemes, in a small-numbers setting.

2. Tournaments versus Termination Schemes

(A) Tournament Schemes and Collusion

The following discrete-time scenario, incorporating a performance-monitoring structure that is in the spirit of that employed in Shapiro and Stiglitz [1984], constitutes the focus of our enquiry in this section. Agents are assumed to have only two feasible effort-choice levels, $\mu \in \{0,1\}$ where $\mu = 0$ represents "shirking." The monitoring structure generates the information $\{\tilde{I}_{\dot{J}}\}$ for agents putting in efforts $\{\mu_{\dot{J}}\}$:

$$\mu_{\dot{J}} = 1 \Rightarrow \tilde{I}_{\dot{J}} = 1 \quad \text{with probability } 1 \tag{1a}$$

$$\mu_{\dot{J}} = 0 \Rightarrow \tilde{I}_{\dot{J}} = \begin{cases} 0 & \text{with probability } P \\ 1 & \text{with probability } (1 - P) \end{cases} \tag{1b}$$

where $P \in (0,1)$ measures the exogenous effectiveness of monitoring. Conditional on $\{\mu_{\dot{J}}\}$, the monitoring noises $\{\tilde{I}_{\dot{J}}\}$ are assumed to be independent across K agents, who choose their effort levels to maximize:

$$\underset{\mu_{\dot{J}}}{\text{Max}}\ Z_{\dot{J}} \equiv \underset{\{\tilde{I}_{\dot{J}}\}}{E} \left[U(\tilde{W}_{\dot{J}}) - C(\mu_{\dot{J}})\right] \tag{2}$$

where $\tilde{W}_{\dot{J}}$ is agent $\dot{J}$'s wealth which is a (measurable) function of $\{\tilde{I}_{\dot{J}}\}_{\dot{J}=1}^{K}$, $U(\cdot)$ is the common concave utility function with $U(0) = 0$, $C(\cdot)$ is the effort-disutility function, normalized at $C(0) = 0$ and $C(1) = 1$, $E(\cdot)$ denotes expectations, and tildes "~" distinguish random variables from their realizations.

Shirking by a worker leads to zero marginal product for him/her, whereas $\mu_{\dot{J}} = 1$ leads to the marginal product $F'(\ell)$, where ℓ is the level of employment and $F(\cdot)$ is the production function (given capital stock), and primes denote derivatives. Given N potential workers per representative firm, we denote the full employment marginal product as:

$$M \equiv F'(N) \tag{3}$$

It is assumed that $F'(\cdot) > 0$, $F''(\cdot) \leqslant 0$, and that $[U(M) - U(0)] > C(1) = 1$, so that allocational (Pareto) efficiency calls for no shirking and full employment, i.e., $\ell = N$ at each firm.[7]

In an environment without principal's moral hazard, the following very simple reward structure--termed a "forcing contract," based on the "moving support" nature of $\{I_{\dot{J}}\}$ in (1a, b)--achieves the first-best (competitive) allocation, which elicits effort without imposing any risk on the risk-averse workers in equilibrium. This extreme risk-sharing goal is optimal given either risk-neutral employers, or a large (countably infinite) N, and satisfying it requires that the following threat be credible. Let the wage for worker $\dot{J}$ be given as;

$$\begin{aligned} \tilde{W}_{\dot{J}} &= M \quad \text{if} \quad I_{\dot{J}} = 1 \\ &= 0 \quad \text{if} \quad I_{\dot{J}} = 0 \end{aligned} \tag{4}$$

by assumption, negative wages are not feasible. It is easy to check that, given full employment and the wage rule (4), a worker maximizing as in (2) will set $\mu_{\dot{J}} = 1$ if

$$P\,U(M) \geqslant C(1) = 1 \tag{5}$$

i.e., the expected utility penalty from shirking exceeds the utility of leisure thus obtained. We assume, in what follows, that condition (5) is satisfied.

Consider now a scenario with two-sided moral hazard, in which only the principal observes the set $\{\tilde{I}_{\dot{J}}\}$, or at least that third-party verification of such indicators is not feasible. Clearly, contracts of the form of equation (4) are not incentive-compatible for such a principal, since he can lie and say that $\{\tilde{I}_{\dot{J}}\} = 0$ for all (many) $\dot{J}$.[8] As discussed in the introduction, one resolution to this problem is to reward agents using a tournament. In the present context, that results in a payoff structure of the form $\tilde{W}_{\dot{J}} = H$ if Rank $(\dot{J}) = 1$, and $\tilde{W}_{\dot{J}} = L$ if Rank $(\dot{J}) = 0$, where Rank $(\dot{J})$ is the realization of a map from the set of $\{I_{\dot{J}}\}$ to lotteries on $\{0,1\}$ that are (conditionally) independent across agents. These lotteries do, however, have the property that for large N the proportion of agents receiving L equals a prespecified Π (almost surely), so that the principal's total wage bill always equals $[(1-\Pi)H + \Pi L]$, as detailed below.

To obtain incentive effects from such a scheme, assume further that agents are told that a $\dot{J}$-th agent who is caught shirking will have first priority in being put into the Rank $(\dot{J}) = 0$ category. Specifically, if the fraction of agents who are detected to be shirking equals S, then the ranking/payoff rule is:

$$\text{Rank}(\dot{J}) = 0 \quad \text{if } I_{\dot{J}} = 0, \quad S \leq \Pi \tag{6a}$$

$$= 1 \quad \text{if } I_{\dot{J}} = 1, \quad S \geq \Pi \tag{6b}$$

$$= 0 \text{ with probability } \frac{\Pi}{S} \text{ if } I_{\dot{J}} = 0, \quad S > \Pi \tag{6c}$$

$$= 0 \text{ with probability } \frac{(\Pi-S)}{(1-S)} \text{ if } I_{\dot{J}} = 1, \quad S < \Pi \tag{6d}$$

with the obvious complementary probabilities. Agents' reactions to such a reward scheme must be analyzed as Nash equilibria of the subgame generated by the scheme, and the elimination of Nash equilibria which are better for agents than that preferred by the principal constitutes a part of the goal in the design of such rewards.

Consider, first, the requirements for $\{\mu_{\dot{J}}\} = 1$, for all $\dot{J}$, to be a Nash equilibrium, which is that:

$$\left[(1-\Pi)\ U(H) + \Pi\ U(L)\right] - 1 \geqslant \left[(1-P)(1-\Pi)\ U(H) + \left\{(1-P)\Pi + P\right\} U(L)\right] \tag{7a}$$

or, equivalently,

$$P(1-\Pi)\left[U(H) - U(L)\right] \geqslant 1 \tag{7b}$$

In establishing (7a) we have used (6a-d) and (1a, b) and observed that, given others are setting $\mu = 1$, the $\dot{J}$-th agent who also works has probability $(1-\Pi)$ of suffering the wage penalty, whereas if he shirks, then with probability P he is so detected and thus obtains wage L for sure, and with the complementary probability he is <u>not</u> detected and thus suffers "randomization among equals."

If equation (7b) were sufficient to ensure incentive compatibility, then with N large we would approach the performance of the first-best forcing contract arbitrarily closely, by setting $\Pi = 1/N$, $L = 0$, $U(H) \geqslant \frac{1}{P(1-1/N)}$, which would be feasible given the (competitive) principal's <u>resource-balance constraint</u>,

$$\left[(1-\Pi)H + \Pi\ L\right] = M \tag{8}$$

for sufficiently large N if equation (5) is satisfied with (infinitesimal) slack. However, the problem of contract design here is not so simple. With such an infinitesimally small Π, agents would be tempted to "collude," e.g., $\{\mu_{\dot{J}}\} = 0$ for all $\dot{J}$ may be a Nash equilibrium and if it is then it would strictly Pareto-dominate the other one above for agents, since they now enjoy leisure and yet (by symmetry) have the same wage distribution, on the equilibrium path. We now examine how large Π has to be to ensure that $\{\mu_{\dot{J}}\} = 0$ is not a Nash equilibrium.

If other agents are setting $\{\mu_{\dot{J}}\} = 0$, then a $\dot{J}$-th agent strictly prefers to set $\mu_{\dot{J}} = 1$ if $0 < \Pi \leqslant P$ (sensibly) and high enough so that:

$$U(H) - 1 > (1-\Pi)\, U(H) + \Pi\, U(L) \qquad (9a)$$

or,

$$\Pi\left[U(H) - U(L)\right] > 1 \qquad (9b)$$

In deriving (9a), we have again used (6a–d) and (1a, b) and noted that now setting $\mu_{\dot{J}} = 1$ results in $W_{\dot{J}} = H$ with probability one, whereas setting $\mu_{\dot{J}} = 0$ results in probability P of getting caught, followed by conditional probability $(\frac{\Pi}{P})$ of punishment, since the detected fraction of shirkers S equals P almost surely. Note that weak preference for $\mu_{\dot{J}} = 1$ is no longer enough, since the $\{\mu_{\dot{J}}\} = 0$ Nash equilibrium, if it exists, is strictly better for agents. It is also easy to show, using (6a–d), that if equations (7b) and (9b) as well as (8) hold, then there are no symmetric (mixed strategy) Nash equilibria other than $\{\mu_{\dot{J}}\} = 1$. This follows because the impact of effort on reducing the probability of "being caught shirking" is the least when all other agents are shirking. We have thus established:

Proposition 1: With two-sided moral hazard, the tournament contract $\{H, L, \Pi\}$, satisfying the "priority rule" of equations (6a-d), obtains the $\{\mu_{\dot{J}} = 1\}$ "work, work" Nash equilibrium uniquely if and only if equations (7b), (8), and (9b) are satisfied (given competitive firms).

For risk-neutral workers, there is a very interesting corollary to the effect that, despite the requirement of eliminating the collusive equilibrium, tournament contracts can achieve the first-best with full employment, whenever individualistic forcing contracts are capable of doing so, as noted in the following result.

Proposition 2: If $U(W) = W$, and equation (5) is satisfied, then there exist $\{H, L, \Pi\}$ which satisfy equations $\{(7b), (8), (9b)\}$, and result in $E[U(\tilde{W}_{\dot{J}})] = U(M)$ for all $\dot{J}$, i.e., first-best welfare levels.

Proof: Pick $L = 0$, $H = (M + 1 + \varepsilon)$, $\Pi = \frac{(1+\varepsilon)}{H}$ for $\varepsilon > 0$ and arbitrarily small. Clearly (9b) is satisfied since

$$\Pi(H - L) = (1 + \varepsilon) > 1$$

and equation (8), for competitive wage payment, is also met because

$$(1-\Pi)H = (H - 1 - \varepsilon) = M$$

and (7b) is met given that

$$P(1-\Pi)(H - L) = PM \geq 1$$

by equation (5). Q.E.D.

However, the implication of the randomization in wages, brought about by a nontrivially large Π that is unresponsive to (any large) N, implies that tournament schemes will not achieve the first-best for strictly risk-averse workers, and may not even be able to implement[9] the (nontrivial) $\{\mu_{\dot{J}}\} = 1$ Nash equilibrium at full employment unless there is a sufficiently large slack in equation (5). We consider these issues, as well as the comparison with the alternative of termination schemes, in what follows.

(B) Termination Schemes and Welfare Comparisons

As Shaprio and Stiglitz [1984] have noted, an alternative solution to the principal's moral hazard problem arises in an intertemporal setting, where workers live for an (unbounded) infinite number of periods. The threat of being terminated from employment in future if detected to be shirking $(I_{\dot{J}} = 0)$ provides the requisite incentive for effort. In the context of our discrete-time model, this idea may be captured by having infinitely-lived workers with preferences given by

$$Z_{\dot{J}} = E \sum_{t=1}^{\infty} \beta^t [U(\tilde{W}_{\dot{J}t}) - C(\mu_{\dot{J}t})], \qquad (0 < \beta < 1) \tag{2'}$$

It is also assumed that the wage-level W conditional on employment, given the employment level ℓ (per firm) satisfies the competitive constraint:

$$W = F'(\ell) \geqslant M \tag{10}$$

Thus, since the wage level per se is unaffected by the principal's report of performance, there is no termination except for incentives purposes.

If a worker is caught shirking, with probability P in any period t in which he sets $\mu_t = 0$, then he remains unemployed for a random length of time $\tilde{T}$, during which he receives the (net) utility level of zero, i.e., $[U(0) - C(0)]$. It is clear that the "expected discounted unemployment spell"

$$Y = \underset{\tilde{T}}{E} \sum_{t=1}^{\tilde{T}} \beta^t \tag{11}$$

is finite. Given our monitoring structure, and the discrete-period modeling, $\tilde{T}$ <u>may</u> even be infinite with probability one, i.e., agents who shirk (off the equilibrium path) would be unemployed forever, whereas those who are unlucky enough to be initially unemployed remain in that state forever. In contrast, Shapiro and Stiglitz's continuous time model requires far more careful endogenization of $\tilde{T}$, given exogenous assumptions about "quit rates" and thus reabsorption, which we do not pursue here.[10]

With the monitoring structure of (1a, b) in each period, it is easy to see that an agent works in a given period if:

$$Y\, P\, [U(W) - 1] \geq 1 \tag{12}$$

i.e., the expected penalty from losing employment given shirking exceeds the utility of leisure. In line with Shapiro and Stiglitz, we shall assume that this incentive condition is <u>not</u> satisfied at the full employment level of $\ell = N$, i.e.,

$$Y\, P\, [U(M) - 1] < 1 \tag{13}$$

However, given <u>strict</u> concavity of the production function $F(\cdot)$, it is assumed that there exists an employment level ℓ such that the competitive wage level, satisfying equation (10), fulfills the incentive condition (12). Note, however, that such an incentive scheme has a clear "deadweight" cost (even for risk-neutral workers), that of workers being "involuntarily" unemployed, i.e., $U(M) > 1$. Notice also that even if $Y > 1$, inequality (13) may be met despite (5) being satisfied, i.e., full employment is consistent with incentives given (unimplementable) individualistic forcing contracts. The reason is that it is only the threat of losing the net surplus from employment in future that provides incentives for effort in a termination scheme, whereas it is the threatened loss of the whole current-period wage that disciplines workers in a contract like that of equation (4) above.

Despite the simplifications resorted to above (relative to the Shapiro-Stiglitz model), the comparison of welfare levels attainable from termination versus tournament schemes is far from straightforward. It is also difficult to predict if termination-based schemes would necessarily arise, even if their welfare levels exceed those obtainable from tournament schemes, i.e., the (ex ante utility) cost of "involuntary" unemployment is less than that due to the ex post randomization in incentive schemes, when each of these schemes is considered in isolation. The reason is that, in an equilibrium across firms, a subset of firms operating on tournament schemes exercises a negative externality on the functioning of termination schemes at other firms, because the "pool of unemployed" (and thus Y in a more completely specified model) is thereby reduced. However, given Proposition 2 above, it is straightforward to establish the following result, which is dependent on the further assumption that on average (ex ante) full employment is better for workers, i.e.,

$$\frac{d}{d\ell}\left[\ell\, U\big(F'(\ell)\big) - \ell\right] > 0 \tag{14}$$

Proposition 3: If workers are risk-neutral and equation (5) is satisfied, i.e., atemporal (moving support -based) forcing contracts can implement incentives for setting $\mu = 1$, then tournament schemes with full employment (i) strictly dominate termination schemes in the attainable ex ante (prior to knowledge of employment) welfare level; and (ii) necessarily come about as the only incentive scheme in competitive equilibrium, given assumptions embodied in equations {(1a-b), (2'), (6a-d), (13), (14)} above.

Sketch of Proof: Proposition 2, together with the assumption in equation (14), establishes dominance since ex ante workers obtain $\left\{\frac{\ell}{N}\,[F'(\ell) - 1]\right\}$ in a termination scheme. The positive prediction follows from the fact that any (hypothetically) unemployed worker can be attracted by a tournament scheme, which provides incentives as well as the first-best net utility level $[F'(\ell) - 1] > 0$ per period, which is the same as that attained for nonshirkers in termination schemes. Thus termination schemes do not provide any incentives for effort.[11] Q.E.D.

We conclude this section with an example, in which strict risk-aversion on the part of workers implies that a tournament scheme which may generate incentives for effort at full employment nevertheless results in a net (expected) utility level of zero. In contrast, employed workers in a termination scheme environment do obtain some net surplus in employment, whereas unemployed workers obtain zero. Thus, prior to the first-time (and perhaps everlasting) selection of those who are chosen to be employed, the ex ante welfare level from a termination scheme strictly exceeds that from a full-employment tournament scheme. This is far from a complete analysis, however, as the arguments above have suggested. If firms have the choice to offer tournament schemes to the unemployed (off the equilibrium path), then the postive analysis of termination schemes must be significantly modified and extended. A "compelling" argument for termination schemes prevailing can only be made if, given the level of unemployment that sustains them and assuming, unrealistically, representative firms, the net expected utility from a tournament scheme is strictly negative, so that a worker prefers unemployment to participation in such a scheme.

Example: Consider a not-too-risk-averse untenured academic worker with utility $U(W) = \sqrt{W}$, for whom (summer) research support provides the only utility beyond subsistence $(U = 0)$ level, and this is denied in the events of (a) being caught shirking on research, or (b) being told he is shirking, as a result of random ("political") events which are allowed to penalize Π fraction of workers! Thus $L = 0$, $U(L) = 0$, $H = \frac{M}{(1-\Pi)}$ and $\Pi = \frac{1}{\sqrt{H}} \equiv \frac{1}{h}$ to satisfy (8) and (9b), the latter only weakly, implying from the product $H(1-\Pi) = M$ that $[h^2 - h - M] = 0$, or that

$$h = \frac{1}{2} + \sqrt{M + \frac{1}{4}} \tag{15a}$$

Now, for condition (7b) to be met, it must be that $P(1-\Pi)\,h \geqslant 1$, or

$$h - 1 \geqslant \frac{1}{P} \tag{15b}$$

which on using (15a) implies that $M \geqslant [\frac{1}{P^2} + \frac{1}{P}]$, or for $U(M) = \sqrt{M}$

$$P\ U(M) \geqslant \sqrt{1 + P} \tag{15c}$$

Notice that the slack required in (5) for (15c) to be satisfied is increasing in P, somewhat surprisingly.

Our worker's expected utility in employment is given by $Z_T = [(1-\Pi)U(H) - 1]$, or $(h - 2)$, and thus using (15a) we have,

$$Z_T = \sqrt{M + \frac{1}{4}} - \frac{3}{2} \tag{15d}$$

which contrasts with the first-best level Z_F (obtainable with individualistic forcing contracts, i.e., "no politics") given by

$$Z_F = \sqrt{M} - 1 \tag{15e}$$

If the "surplus in employment" $(\sqrt{M} - 1)$ is, for example, equal to $(\sqrt{2} - 1) \sim .432$, for $M = 2$, then the tournament scheme generates (net) expected utility $Z_T = \sqrt{2.25} - 1.5 = 0$. This may seem quite damaging, but for the observation that the associated Π equals $\frac{1}{h} = .5$, using (15a), i.e., this is quite a political employer! Notice that (5) is satisfied as long as $P \geqslant \frac{1}{1.432} \sim .716$, but (15c) is satisfied only if $P = 1$. Notice also that the termination-based scheme is viable with full employment, i.e., (13) is reversed, with $P = .716$ if $Y \geqslant 3.53$ and with $P = 1$ if $Y > 2.32$. For lower values of (possibly endogenized) Y, a parametric form for $F(\ell)$ is needed, in order to endogenize the employment level ℓ that would sustain a termination-based scheme.

In deriving the above results, and calculating the example, we have ignored the role of other sources of ex post information that may be common knowledge, like individual or aggregate (firm-level) output. Even if the former is not observable, observability of aggregate output is going to limit the principal's ability to misrepresent $\{\tilde{I}_j\}$, i.e., to penalize a large fraction of workers as being shirkers even though the "work, work" Nash equilibrium is obtained. This would loosen the relationship between the principal's "power to penalize" (Π) and equilibrium penalties, and thus improve the performance of tournament schemes for strictly risk-averse workers. This topic also merits further research, before a substantial judgment regarding the comparison with termination schemes can be reached.

In the next section we consider alternatives to tournament schemes, and their relative performances, when (strictly risk-averse) agents' utility functions are unbounded below, so that the threat of rank-dependent penalties is able to obtain (nearly) first-best allocations somewhat more robustly.

3. Tournaments and Asymptotic Forcing Contracts

As we noted in the Introduction, much of the earlier rationalization of tournament contracts, e.g., Lazear and Rosen [1981] and Bhattacharya [1982], was based on the observation that comparisons across agents provided a means of controlling for correlated shocks affecting the productivity of (all) both agents. In Lazear and Rosen, as well as in Nalebuff and Stiglitz [1983], this observation has been followed by comparing the performance (in inducing effort and expected utility levels) of tournaments with that of restricted linear piece-rates. As Holmström [1982] has noted, without such an ad hoc restriction on the alternative, tournaments are likely to be dominated by more general multiperson incentive schemes tied to individual performance and a (Blackwell [1951]) "sufficient statistic" for the common (correlated) shock. Green and Stokey [1983] have shown, in a very special case where the common shock does <u>not</u> affect the <u>marginal</u> productivity of effort, that a sequence of tournament schemes with an increasing number of agents and ranks approaches the second-best optimal incentive contract, i.e., the set of order statistics is sufficient with respect to the effort levels.

Our interest in such comparisons is qualified by the following observations. In the small number of agents case, the restriction of the comparison set to linear contracts is arbitrary, and it also violates the coarse monitoring notions implicit in rank-order comparisons. It is an interesting observation that with principal's moral hazard also present and a large number of agents, incentive-compatible contracts in which rewards are a linear function of differences between individual and <u>group mean</u> products converge (with independent shocks) to linear piece-rates. However, as Green and Stokey have noted, in some (other) circumstances tournaments (nonlinear contracts) are able to do much better than such linear schemes. Nor are we persuaded that the recent results of Holmström and Milgrom [1985], on the optimality of linear contracts in an intertemporally stationary continuous-time model with exponential utility and Normal error distributions, hold with any generality.[12]

An interesting and incremental observation about tournaments has been made in another result of Nalebuff and Stiglitz [1983], for the case where agents' utility functions are unbounded below, for feasible payoffs. They show that given some conditions on convolutions of the distribution function for monitoring noises, and a large number of agents, asymptotically a contract based on punishing only the worst-performing agent approaches the performance of the first-best optimal allocation with perfect information. Since this result (unlike that of Green-Stokey) allows for the effect of correlated shocks on the marginal product of effort, it is of interest to ask if other nontournament schemes have such performance under similar circumstances.

A clear candidate for such an alternative scheme is that due to Mirrlees [1979], which is based on individualistic contracts (without correlated shocks), as well as unbounded utility and "likelihood ratio" assumptions about preferences and output (monitoring) noises, respectively, as the wealth/output level becomes as low as feasible. We show in the Appendix that the Nalebuff-Stiglitz distributional assumptions imply those made by Mirrlees, and that the latter's incentive-scheme is extendable to the correlated shock environment of Nalebuff-Stiglitz, so that both schemes (which are not equivalent in ex post payoffs) approach first-best optimality as the number of competing agents becomes large. In this section, we introduce the notational and analytical constructs involved in obtaining these results (Theorem 1 and Proposition 4 in the Appendix), which pertain to (i) distributional conditions in the lower tail[13] of monitoring errors, and (ii) sufficient statistics for the correlated shock gathered from individual performance indicators. Note that two-sided moral hazard is being assumed away here, since the Mirrlees incentive scheme is not incentive-compatible for the principal in such a scenario.

In Nalebuff and Stiglitz [1983], it is assumed that the observed output/monitor of the $\dot{J}^{th}$ agent $\tilde{d}_{\dot{J}}$ satisfies the equation,

$$\tilde{d}_{\dot{J}} = \tilde{\theta}\mu_{\dot{J}} + \tilde{e}_{\dot{J}} \tag{16}$$

where $\mu_{\dot{J}}$ is his/her effort level, $\tilde{\theta}$ is a <u>common</u> shock, and $\{\tilde{e}_{\dot{J}}\}$ are identically distributed shocks satisfying statistical independence with respect to $\tilde{\theta}$ and $\{\tilde{e}_i\}$, $i \neq \dot{J}$. Agents are assumed to be risk-averse, and to maximize:

$$\underset{\{\mu_{\dot{J}}(\theta) \geq 0\}}{\text{Max}} \; Z_{\dot{J}} \equiv \underset{[\tilde{\theta}, \{\tilde{e}_{\dot{J}}\}]}{E} \left[U(\tilde{W}_{\dot{J}}) - C(\mu_{\dot{J}}(\tilde{\theta}))\right] \tag{2''}$$

where $\tilde{W}_{\dot{J}}$ is a (measurable) function of $\{\tilde{d}_{\dot{J}}\}_{J=1}^{K}$, K is the number of agents, and effort levels $\mu_{\dot{J}}(\theta)$ are chosen <u>after</u> agents (only) observe the realized common shock θ. They show that the first-best optimal allocation, obtainable with competitive risk-neutral employers who may observe and dictate employees' effort levels, satisfies the conditions:

$$\tilde{W}_{\dot{J}} = Y \qquad \forall \; \theta, \; \{e_{\dot{J}}\} \tag{17a}$$

$$\theta \, U'(Y) = C'(\mu^{*}(\theta)) \quad \forall \; \theta \tag{17b}$$

$$Y = \underset{\tilde{\theta}}{E} \left[\tilde{\theta} \, \mu^{*}(\tilde{\theta})\right] \tag{17c}$$

Equations (17a–c) can be interpreted as the employer providing complete insurance, and eliciting effort to the point where its marginal utility product equals its marginal disutility.

The following "punish the worst" incentive contract approaches the allocation in (17a-c) arbitrarily closely, as K (the number of agents) becomes arbitrarily large. After ranking by decreasing output from $\{1,2,\ldots,K\}$, the j^{th} ranked agent is paid;

$$W_j = Y + X(K) \quad \text{for} \quad j = 1,2,\ldots,(K-1) \tag{18a}$$

$$W_K = Y - (K-1)\; X(K) \tag{18b}$$

where the penalty factor $X(K)$ satisfies, given

$$\Delta U_K \equiv \left[U(W_1) - U(W_K)\right] \tag{18c}$$

the agents' (first-order) incentive condition;

$$\theta\; H(K)\,\Delta U_K = C'\left(\mu^*(\theta)\right) \;\forall\; \theta \tag{19a}$$

which, on using (17b), simplifies to

$$H(K)\,\Delta U_K = U'(Y) \tag{19b}$$

where $\left[\theta\; H(K)\right]$ is the equilibrium marginal impact of effort μ on the probability of finishing last, derived as follows.

Assuming that other agents are putting in effort $\mu^*(\theta)$ in state θ, i.e., analyzing one Nash equilibrium, and given the specification of equation (16), an agent putting in effort μ calculates the probability $P(\mu, K)$ of coming in last as:

$$P(\mu, K) = \int_{-\infty}^{\infty} \{1 - G[e + \theta(\mu - \mu^*(\theta))]\}^{K-1} g(e)de \qquad (20a)$$

where $\tilde{e}$ is <u>generic</u> for the random variables $\{\tilde{e}_j\}$ which are <u>i.i.d.</u>, and $\{G(e), g(e)\}$ represent its distribution/density functions. For simplicity, $\tilde{\theta}$ is assumed to have a distribution with bounded support, but the unbounded (below) support of $G(e)$ rules out simple individualistic forcing contracts based on moving (with μ) supports. Given (20a), the equilibrium marginal effect of effort on "finishing last" is given by:

$$\theta\ H(K) \equiv - \frac{\partial P(\mu,K)}{\partial \mu} \Big|_{\mu=\mu^*(\theta)} = \theta(K-1) \int_{-\infty}^{\infty} [1 - G(e)]^{K-2} g^2(e)de. \qquad (20b)$$

The Nalebuff-Stiglitz "approaching the first-best" argument (as $K \to \infty$) is then based on the observation that even though (it may be that)

$$\underset{K\to\infty}{\text{Limit}}\ H(K) = 0 \qquad (21a)$$

given that range of $\tilde{e}$ is unbounded below, for some distributions it is the case that

$$\underset{K\to\infty}{\text{Limit}}\ K\ H(K) = \infty \qquad (21b)$$

Hence, $[\frac{\Delta U_K}{K}]$ which is the (symmetric) equilibrium "expected utility loss" approaches zero as $K \to \infty$ while satisfying the incentive equation (19b) for each K. However, given (21a), it must be the case that $\Delta U_K \to \infty$ as $K \to \infty$, given (19a) which is the first-order condition to the agents' maximization problem (2''), given (18a, b). Other Nash equilibria are ruled out given strict convexity of $C(\mu)$, and the unbounded utility penalty--which substitutes (much more effectively) for the "number of punishable agents" in the model of section 2 above.

In the Appendix we show that (21b) implies the Mirrlees [1979] "unbounded likelihood ratio" condition, which in the context of the specification (16) is equivalent to

$$\underset{e \to -\infty}{\text{Limit}} \frac{g(e)}{G(e)} = \infty \tag{22}$$

assuming only some innocuous regularity (differentiability and boundedness) assumptions about $\{G(e), g(e)\}$. This purely statistical result, which (unfortunately) requires a lengthy proof, constitutes Theorem 1. Following this we show that the first-best allocation in (17a-c) may be arbitrarily closely approached by the alternative of an extended Mirrlees contract, which works as follows.

Agents' observed productivities $\{d_{\dot{J}}\}$ are collected, and a "sufficient statistic" for θ vis-à-vis the $\dot{J}^{th}$ agent is calculated using $S_{\dot{J}} = \sum_{i \neq \dot{J}} d_i$, which is then inverted using the $\{\theta\ \mu^*(\theta)\}$ mapping to obtain the "guessed" $\hat{\theta}_{\dot{J}}$. This guess is then used, together with incentive conditions of the form of (19a), to penalize agent $\dot{J}$ with progessively higher penalties for outcomes sufficiently far in the lower tail. The appropriate double-limit of this incentive-constrained allocation, with respect to the "punishment threshold" and the number of agents, is shown in Proposition 4 (Appendix) to approximate the first-best allocation arbitrarily closely. Thus, the inessentiality of tournament (or other constant-sum-across-agents payoff) schemes, in environments without principal's moral hazard, is further established.

4. Concluding Remarks

We have shown that multiagent contracts with specific ex post properties (of being constant-sum across agents), may have a significant role to play even in the absence of correlated outputs or measurements of performance, if there is moral hazard on both sides in the implementation of allocations. Tournament contracts, which provide a specific example of such schemes, were analyzed and their performance vis-à-vis alternative individualistic (termination) schemes was shown to be highly dependent, given the problem of "collusion" across agents, on (a) agents' risk-aversion levels, and (b) the availability of aggregate (as well as individual) measures of productivity. The degree to which the interaction between principal's moral hazard and agents' collusion problems is (empirically) important, as well as the general equilibrium endogenization of multiagent versus individualistic schemes in such circumstances, remain open as significant topics for further enquiry. Only the basic issues and tradeoffs, including "externalities" across multiple principals (firms), have been developed here.

We also showed that, in the large number of agents case, the constant-sum-across-agents property of tournament schemes is inessential to the achievement of allocations they have been "identified" with, with correlated shocks but without principal's moral hazard. In an important sense, however, the focus on such asymptotic results--here as well as in Green-Stokey [1983] and Nalebuff-Stiglitz [1983]--misses an important set of issues that were initially sought to be addressed by the early work of Lazear and Rosen [1981]. In essence, these issues pertain to explaining the prevalence of such multiagent (ordinal) comparison schemes in hierarchical organizations, largely confined to contests within levels of the hierarchy, and implemented despite the presence of a small number of agents in each contest.

It is our belief that despite much research effort, a great deal of success has not been achieved in explaining such contracts, vis-à-vis the alternative of individualistic contracts with other (firm-or industry-wide) measures used to adjust for correlated shocks. Within such hierarchies, it is also important to take into account differences in productive abilities across agents, which may be asymmetrically known and in need of elicitation through contracts. Further exploration of some of these issues, including corrections and extensions of earlier (negative) results, e.g., Lazear and Rosen [1981], is to be found in Bhattacharya and Guasch [1985].

FOOTNOTES

[1]Lazear and Rosen [1981] were the first to note this explicitly, although the main focus of their analysis is different. Holmström [1983] provides a general result regarding such linkage in incentive schemes that is related to the Blackwell [1951] notion of (non)sufficient statistics.

[2]Bhattacharya [1983] and Malcolmson [1984] independently made this observation in the context of atemporal multiagent incentive schemes, and the former was the first to notice the complexities arising from potential "collusion" across agents, discussed below.

[3]In Shapiro and Stiglitz [1984], which is a continuous-time model, there is an exogenously specified quit rate of workers so that unemployed workers have a stationary probability of being reemployed, i.e., a moment's indiscretion (shirking) does not cause infinite misery!

[4]Other authors have pointed out that termination schemes may arise even in the absence of two-sided moral hazard, given restrictions on contingent contracts but with sufficient intertermporal linkages to make resulting payoffs sustainable.

[5]Since principal's moral hazard prevents individualistic output/monitoring-contingent reward schemes, the incentive problem is "nontrivial" even for risk-neutral workers.

[6]The extended Mirrlees incentive scheme does not have the property that payoffs summed across agents are constant (with probability one). We also do not pursue the topic of comparing tournaments with arbitrarily restricted, e.g. linear, payoff schemes. See below for further details on these issues.

[7]The representative firm assumption is innocuous, in that $F(\cdot)$ can instead be an aggregate production function which is obtained by allocating labor efficiently across firms. The bounded-below utility function assumption, i.e., $U(0) = 0$, is important for the results to follow.

[8]Of course, if an auxiliary indicator like aggregate output is also observable, then that limits (subject to the noise in output given efforts) the extent of such lying. We do not explicitly consider this complication here although, given the problem with tournament schemes that are identified below, the "disciplining" value of such aggregate indicators is important; see below.

[9]While Proposition 2 is intuitively pleasing, it should be noted that it arises due to the colinearity between equations (7b) and (8) given $U(W) = W$, rather than any general game-theoretic notions.

[10]The essence of this difference arises because infinite spells of unemployment act essentially like unbounded utility penalties, which can result in attaining the first-best allocation far more easily; see section 3 below.

[11]The representative-firm notation regarding the production function is clearly overextended in an argument of this sort. For contreteness, envision the tournament offerers to be "renegade" personnel officers in other (divisions of) similar firms.

[12]The essence of this result arises from the observations that the Normal distribution is infinitely divisible into binomials, over which (instantaneous) contracts are necessarily linear, together with the constant absolute risk-aversion properties of exponential utility functions, i.e., piece-wise linear contracts with the same slopes add up to linearity globally. A helpful analogy may be obtained by noting that for the Black-Scholes [1973] option-pricing model, developed with similar probabilistic assumptions, the option-stock "hedge ratio" is constant if the exercise price changes in proportion to stock price continuously. (Variance is exogenously given in both cases.)

[13]It is easy to show that analogous results, on the relation between convolutions and likelihood ratios of distribution functions, also hold for the upper tail. This is not pursued, since as Nalebuff-Stiglitz [1983] have noted, incentive schemes based on unbounded (utility) bonuses for the best-performing agent fail to generate appropriate incentives when the number of agents becomes large. Thus, by implication, the attempted extension of Mirrlees' result in Holmström [1982] Theorem 4 is incorrect; agents minimize "incentive compatibly" in such schemes.

APPENDIX

In what follows, we examine the condition on (convolutions of) the probability distribution function $G(e)$ which, together with utility functions $U(\cdot)$ that are unbounded below, imply that asymptotically (as the number of agents K approaches infinity) the first-best effort and expected utility levels are approached arbitrarily closely. This condition, developed in Nalebuff and Stiglitz [1983] and discussed in section 3 above, can be expressed as:

$$\underset{K\to\infty}{\text{Limit}}\ [K\ H\ (K)] \equiv K(K-1) \int_{-\infty}^{\infty} [1 - G(e)]^{(K-2)}\ g^2(e)de = \infty \qquad \text{(A1)}$$

where $g(\cdot)$ is the density of $G(\cdot)$, and $\theta\ H(K)$ has the economic interpretation of being the marginal effect of effort level μ on the probability of "coming in last" in the tournament, which is denoted in the text as $P(\mu, K)$.

We first show that condition (A1) implies, given some regularity conditions on $G(\cdot)$, that

$$\underset{e\to -\infty}{\text{Limit}} \frac{g(e)}{G(e)} = \underset{e\to -\infty}{\text{Limit}} \frac{g'(e)}{g(e)} = \infty \qquad \text{(A2)}$$

where primes denote derivatives, and the middle equality follows from L'Hopital's rule. Denoting the conditional output probability distribution function as $\Pi(\mu, d, \theta) \equiv \text{Prob}(\tilde{d} \leq d | \mu, \theta)$, and using the fact that $\tilde{d} = \mu\theta + \tilde{e}$, it is easily seen that (A2) implies that, for all θ and μ,

$$\underset{d \to -\infty}{\text{Limit}} \frac{\partial \Pi / \partial \mu(\mu, d, \theta)}{\Pi(\mu, d, \theta)} = -\infty \qquad (A3)$$

which is the condition used in Mirrlees [1979].

Theorem 1

Assume that $G(e)$ is three-times differentiable, and that there exist finite positive numbers $\{M, N\}$ such that $0 < g(e) < M$, $|g'(e)| < N$, for all e. Then (A1) holds only if (A2) is satisfied.

Remarks: We shall prove the result without assuming that $g'(e)$ is absolutely integrable, i.e., that $\int_{-\infty}^{\infty} |g'(e)|de$ converges, although this property is true of some distributions (like the Normal) which satisfy (A1) and (A2). In the course of the proof, we shall also show that these assumptions imply that $\underset{K \to \infty}{\text{Limit}}\ H(K) = 0$.

To proceed, we need the following result from classical calculus.

Lemma (Abel). For every $K \varepsilon [2, \infty)$, let $C(K, e)$ and $D(K, e)$ as functions of e be such that $C(K, e)$ is continuous and $D(K, e)$ is monotonically decreasing for $e > a$. Assume, furthermore, that there exists a constant M such that $\left|\int_a^b C(K, e)de\right| < M$ for every $K \varepsilon [2, \infty)$ and every $b > a$, and that $\underset{e \to \infty}{\text{Lim}}\ D(K, e) \to 0$ uniformly for all $K \varepsilon [2, \infty)$. Then the integral

$$I(K) = \int_a^{\infty} D(K, e)\ C(K, e)de$$

converges uniformly for $K \varepsilon [2, \infty)$.

Remark: Olmsted [1961], section 1408, notes that uniform convergence as above implies that "the limit of an improper integral (w.r.t. K) is the integral of the limit."

Proof of Theorem 1: Using equation (A1) and integrating by parts, we have

$$\lim_{K\to\infty} H(K) = \lim_{K\to\infty} \left\{ -[1 - G(e)]^{K-1} g(e) \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} [1 - G(e)]^{K-1} g'(e)de \right\} \tag{A4}$$

The first term is zero because $\text{Limit}_{e\to -\infty} g(e) = 0$. Since $g(e) < M$, and $[1 - G(e)]^{K-1}$ is decreasing and converges uniformly (for $K \geqslant 2$) to 0 as $e \to \infty$, the second (integral) term converges uniformly for all K, by the preceding lemma. Hence, we can interchange the order of the limit with respect to K and integration and, using the assumptions that $[1 - G(e)] < 1$ for almost all e, and $|g'(e)| < N$ obtain that $\text{Limit}_{K\to\infty} H(K) = 0$.

To evaluate $\lim_{K\to\infty} [K\,H(K)]$, multiply equation (A4) by K, and split up the integral term into two, over $(-\infty, T\alpha)$ and $[T\alpha, \infty)$, where $T\alpha$ satisfies, for some $0 < \alpha < 1$, $\alpha = [1 - G(T\alpha)]$. Thus, we have

$$\lim_{K\to\infty} K\,H(K) = \lim_{K\to\infty} \left\{ -K[1 - G(e)]^{K-1} g(e) \Big|_{-\infty}^{\infty} + \int_{-\infty}^{T\alpha} K[1 - G(e)]^{K-1} g'(e)de \right.$$

$$\left. + \int_{T\alpha}^{\infty} K[1 - G(e)]^{K-1} g'(e)de \right\} \tag{A5}$$

The first ("nonintegral") term of equation (A5) is again 0, since the inner limit with respect to e is zero for all K. The third term satisfies the conditions of Abel's Lemma, since $K[1 - G(e)]^{K-1}$ is monotone decreasing in e, and converges uniformly in K to 0 as $e \to \infty$.[1] Thus,

$$\lim_{K\to\infty} \int_{T\alpha}^{\infty} K[1 - G(e)]^{K-1} g'(e)de = \int_{T\alpha}^{\infty} \lim_{K\to\infty} K[1 - G(e)]^{K-1} g'(e)de$$

which is zero because $\lim_{K\to\infty} K[1 - G(e)]^{K-1} = 0$ for all $e \in [T\alpha, \infty)$.

Thus, integrating the second term in (A5) by parts, we have

$$\lim_{K\to\infty} K\,H(K) = \lim_{K\to\infty} \left\{-[1 - G(e)]^K \frac{g'(e)}{g(e)} \Big|_{-\infty}^{T\alpha} + \int_{-\infty}^{T\alpha} [1 - G(e)]^K \frac{d}{de}\left[\frac{g'(e)}{g(e)}\right]de\right\} \tag{A6}$$

Now suppose, contrary to the proposition, that $\lim_{K\to\infty} [K\,H(K)] = \infty$ even though (A2) is not satisfied. We show that, for $T\alpha$ sufficiently small in (A6), this cannot be. Consider the first term in (A6), which is then finite under the assumptions of the proposition and the maintained assumption that $\lim_{e\to -\infty} \frac{g'(e)}{g(e)}$ is finite. The second term in (A6) can be broken up as

$$\text{Limit}_{K\to\infty} \left\{ \int_{-\infty}^{T\alpha} \left([1 - G(e)]^K - 1\right) \frac{d}{de}\left[\frac{g'(e)}{g(e)}\right]de + \int_{-\infty}^{T\alpha} \frac{d}{de}\left[\frac{g'(e)}{g(e)}\right]de \right\} \tag{A7}$$

both of which converge uniformly in K, the latter obviously, and the former by Abel's Lemma (with e substituted by $-e$), given the maintained assumption on g'/g, assuming $T\alpha$ sufficiently small. By interchanging the order of the limit with respect to K, we obtain in (A7) the sum of two terms, both of which are finite. Thus, we have established a contradiction to the hypothesis that equation (A1) can hold without equation (A2) also holding. Q.E.D.

We now prove the result which extends the insight of Mirrlees [1979] to the "correlated shock case."

Proposition 4: Assume that $U(W)$ is unbounded below as $W \to 0$, and that (A1) holds, implying by Theorem 1 that (A2) and (A3) are satisfied. For agent j, let the sufficient statistic relating to the common shock $\tilde{\theta}$ be constructed as:

$$\tilde{S}_j \equiv \frac{1}{K-1} \sum_{\substack{i=1 \\ i \neq j}}^{K} \tilde{d}_i \tag{A8}$$

and let, for $m = 1,2,\ldots,\infty$ agent j's performance-contingent payoff sequence be given, for $D_1 > D_2 > \ldots D_m$, by

$$W_j^{Km} = Y \quad \text{for} \quad d_j \geq D_m, \quad W_j^{Km} = Y - P_m \quad \text{for} \quad d_j < D_m \tag{A9}$$

where

$$P_m = Y - U^{-1}\left(U(Y) - B_m\right) \tag{A10}$$

where B_m, the "utility penalty," satisfies

$$B_m = \frac{C'\left(S_j/V(S_j)\right)}{V(S_j)g(D_m - S_j)} \tag{A11}$$

where $V(\cdot)$ is such that

$$V^{-1}(\theta) = \mu^*(\theta)\ \theta \tag{A12}$$

The goal, therefore, is to have the appropriate incentive condition satisfied at $\mu = \mu^*(\hat{\theta})$ for the <u>inferred</u> common state $\hat{\theta}_j = V(S_j)$. Assume further that agent j chooses effort $\mu_j^{Km}(\theta)$ to maximize

$$\underset{\mu}{\text{Max}} \quad Z_j^{Km}(\theta) = E\left[U\left(\tilde{W}_j^{\ Km}\right) \middle| \theta,\ \mu\right] - C(\mu) \tag{A13}$$

assuming that $\tilde{d}_i = \theta\mu^*(\theta) + \tilde{e}_i$, $i \neq j$, i.e., that other agents are putting in effort $\mu^*(\theta)$. Assume also that $g(d - \theta\mu)/C'(\mu)$ is decreasing in μ for all d; this ensures that the agent's objective function is concave in μ. If $\lim_{m\to\infty} D_m = -\infty$ then, for each θ in the range of $\tilde{\theta}$,

$$\lim_{\substack{K\to\infty \\ m\to\infty}} Z_j^{Km}(\theta) = U(Y) - C\left(\mu^*(\theta)\right) \tag{A14}$$

$$\lim_{\substack{K\to\infty \\ m\to\infty}} \mu_j^{Km}(\theta) = \mu^*(\theta) \tag{A15}$$

<u>Remarks</u>: Notice that, given $\lim_{e\to -\infty} g(e) = 0$ and (A11), as $D_m \to -\infty$, the threatened utility penalty $B_m \to \infty$, which is made feasible by $U(\cdot)$ being unbounded below. The nonexistence of equilibria other than the above, in the agents' subgame, was noted in section 3.

Proof: It is well known that, assuming that $\mu_i = \mu^*(\theta)$ for $i \neq j$ and as $K \to \infty$, $\tilde{S}_j \to \theta\mu^*(\theta)$ almost surely. Hence, the agent's objective function (A13) converges uniformly (in μ) to that analyzed in Mirrlees [1979] Theorem 1, from which the result is immediate. Q.E.D.

FOOTNOTES TO APPENDIX

[1] This is true if for any $\delta > 0$ there exists finite $e^*(\delta)$ such that, for all $e \geqslant e^*$ and all $K \in [2, \infty)$,

$$K[1 - G(e)]^{K-1} \leqslant \delta$$

Taking logs, we see that we must have

$$\log K + \log [1 - G(e^*)](K-1) \leqslant \log \delta$$

or,

$$- \log [1 - G(e^*)] \geqslant \frac{\log K}{(K-1)} - \frac{\log \delta}{(K-1)}$$

The claim then follows since $\log (0) = -\infty$, and for $K \geqslant 1$ $(\log K)/(K-1)$ is bounded above by unity, whereas for $\delta < 1$ the second term on the right-hand side is bounded above by $(-\log \delta)$, and otherwise it is negative.

REFERENCES

BHATTACHARYA, S. "Tournaments and Incentives: Heterogeneity and Essentiality." Stanford GSB Working Paper #695 (March 1983).

BHATTACHARYA, S. "Aspects of Monetary and Banking Theory and Moral Hazard." Journal of Finance, Vol. 37 (1982), pp. 371-384.

BHATTACHARYA, S. AND GUASCH, J. L. "Heterogeneity, Hierarchy, and Tournaments." Mimeo 1985.

BLACK, F. AND SCHOLES, M. "The Pricing of Options and Corporate Liabilities." Journal of Political Economy, Vol. 81 (1973), pp. 637-654.

BLACKWELL, D. "Comparisons of Experiments" in J. Neyman, ed., Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, 1951.

CARMICHAEL, L. "Firm-Specific Human Capital and Promotion Ladders." Bell Journal of Economics, Vol. 14 (1983), pp. 251-258.

GREEN, J. AND STOKEY, N. "A Comparison of Tournaments and Contracts." Journal of Political Economy, Vol. 91 (1983), pp. 349-364.

HOLMSTRÖM, B. "Moral Hazard in Teams." Bell Journal of Economics, Vol. 13 (1982), pp. 324-340.

HOLMSTRÖM, B. AND MILGROM, P. "Aggregation and Linearity in the Provision of Intertemporal Incentives." School of Organization and Management, Yale, Working Paper #D5 (1985).

LAZEAR, E. AND ROSEN, S. "Rank-Order Tournaments as Optimal Labor Contracts." Journal of Political Economy, Vol. 89 (1981), pp. 841-864.

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

MIRRLEES, J. A. "The Implications of Moral Hazard for Optimal Insurance." Nuffield College, Oxford (1979).

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

NALEBUFF, B. AND STIGLITZ, J. E. "Prizes and Incentives: Toward a General Theory of Compensation and Competition." Bell Journal of Economics, Vol. 14 (1983), pp. 131-147.

OLMSTED, J. M. "Advanced Calculus." Prentice-Hall Publishing, New York (1961).

SHAPIRO, C. AND STIGLITZ, J. E. "Equilibrium Unemployment as a Worker Disciplining Device." American Economic Review, Vol. 74 (1984), pp. 433-444.

SPENCE, M. AND ZECKHAUSER, R. "Insurance, Information, and Individual Action." American Economic Review, Vol. 61 (1971), pp. 380-387.