Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Multiperson Tournaments: An Experimental Examination

Alannah Orrison, Andrew Schotter, Keith Weigelt,

MANAGEMENT SCIENCE
Vol. 50, No. 2, February 2004, pp. 268–279
ISSN 0025-1909 | EISSN 1526-5501 | 04 | 5002 | 0268

DOI 10.1287/mnsc.1030.0128

Multiperson Tournaments: An Experimental Examination

Alannah Orrison
Department of Economics, Saddleback College, Mission Viejo, California 92692, aorrison@saddleback.edu

Andrew Schotter
Department of Economics, New York University, New York, New York 10003-6687, andrew.schotter@nyu.edu

Keith Weigelt
Department of Management, The Wharton School, 2000 Steinberg Hall—Dietrich Hall, Philadelphia, Pennsylvania 19104-6370, weigelt@wharton.upenn.edu

Modern hierarchical organizations, like corporations, must motivate agents to work hard. Given their pyramid structure, it is not surprising that one commonly used motivator is the promotion tournament. In such tournaments, agents compete to advance to positions at higher organizational levels. Though these tournaments are common, little research has empirically looked at the interface of organizational structure and tournament design. This paper aims to take a step in filling this void by comparing the performance of various tournament designs using controlled laboratory techniques.

Key words: Tournaments; Economic Experiments; Incentives
History: Accepted by Martin Weber, former department editor; received September 24, 1999. This paper was with the authors $3\frac{1}{2}$ years for 1 revision.

1. Introduction

Modern hierarchical organizations, like corporations, must motivate agents to work harder. One commonly used motivating device is the tournament for promotion. In these tournaments, agents compete for promotions that grow in scarcity as they travel up the organizational pyramid. Since only relatively high-performing agents move on to the next organizational level, this tournament also helps assign agents to hierarchical positions.

Organizational structure interfaces with tournament design in the following sense. Agent effort is affected by a set of variables controlled by managers. Several of these are structural—for example, the number of hierarchies, promotions, and tournament size. Structurally, promotion tournaments in firms represent a diverse set, both in breadth and depth. Clearly, across corporations, the variance in the number of positions at any given organizational level is substantial, except at the top level, where most are served by a single chief executive officer (CEO). The number of hierarchical levels also varies across firms.

Intuitively, it seems as if these structural variables should influence agent behavior. For example, one might think having too few promotions leads to discouragement and reduced effort, while having too many promotions leads to agent shirking due to the high inherent probability of getting promoted. Is there an optimal promotion rate? Or, some may believe that organizational growth should be reflected in tournament structure. If true, then managers must structurally change tournaments as the organization grows. Can managers simply replicate the current promotion scheme as the organization grows? Finally, some believe (and have shown) that discrimination reduces the effort of all in promotion tournaments. Might managers use their structural power to design tournaments to buffer these negative effects on agent output? For instance, discrimination may be less damaging in large tournaments than in small ones.

We test these behavioral predictions in a laboratory setting. Though a modest step, we nevertheless believe it is an important one in understanding behaviors in tournaments. Because nonlaboratory research cannot specify tournament parameters, there is little current research regarding behavior in tournaments. This caveat is true even for those studies of well-defined tournaments (e.g., NCAA basketball). We are able to isolate tournaments of specific design and size to study, however. This allows us to measure changes in behavior across the various tournament designs.

We organize our results around three broad themes. These are:

1. *Organizational replication.* The interface of organizational structure and tournament design that managers face is dynamic in nature. Successful firms

grow. How should this growth affect tournament structure? One set of simplifying strategies is replication. In replication, the number of promotions grows proportionally with the number of agents: An organization of 10 agents using a tournament with 5 available promotions would use a tournament with 50 available promotions if they increased in size to 100 agents. One issue, then, is whether agents work harder, the same, or less in the two tournaments.

2. *Prize structure.* A tournament's prize structure motivates agent behavior. What happens to agent effort as the proportion of promotions is changed? For instance, do agents in a 10-agent organization put forth greater effort with 3 promotion possibilities or 6 promotion possibilities?

3. *Organizational size and discrimination.* Discrimination may exist within an organization when the rules differentiate among agents. Favored agents receive preferential treatment, so they are given promotions, though their output may be lower than that of others. Hence, a fraction of agents are discriminated against because their output must be significantly higher than favored agents to receive promotions. Discrimination in tournaments has a negative effect on the output of all agents; firms are always better off without it. Might managers use organizational structure to buffer the negative impact of discrimination?

Tournament theory does shed light on these questions, with some predictions being counterintuitive. While we formally develop this theory in §2, we review some of the main results here. For all of the following results, we assume that agent output is partially determined by a random shock drawn from a uniform distribution with finite support. We note a few results will change if we alter this functional form, though most hold for the entire class of symmetric unimodal distributions, which, of course, includes the normal and uniform.

Somewhat surprisingly, the theory predicts that a tournament's prize structure should not influence agent effort. It predicts that agent effort is homogenous of degree zero for proportionate increases in organizational size. So, as organizational size increases and the proportion of promotions remains constant, expected agent effort is invariant. Hence, at any interior equilibria, a 10-person tournament with 3 promotions should generate the same agent output as a 10-person tournament with 8 promotions. Finally, and perhaps most surprisingly, the theory predicts that size can buffer organizations from discrimination's negative impact on agent output (see Schotter and Weigelt 1992). The efficiency loss of discrimination decreases as tournament size increases. In fact, even in relatively small organizations, agent effort in discriminated tournaments approaches that of fair tournaments.

Interestingly, there is no precise measure of tournament use in organizations. However, a growing body of evidence suggests tournaments are used extensively. Most firms, obviously, have pyramid-shaped organizational structure. As one scales this pyramid, the number of available positions at each new level decreases. Several authors have shown that both differences in compensation and spreads in adjacent levels increase in organizational level (Leonard 1990, Lambert et al. 1993, Bognanno 2001, Conyon et al. 2001). Others have shown that promotion (i.e., change in position) is essential for salary growth (Baker et al. 1993, 1994). There is also evidence that external hiring is more prevalent at lower organizational levels (Baker et al. 1993, 1994; Bognanno 2001).

We summarize some of our main findings here, and provide greater detail in §4. When we replicated tournaments in the laboratory by increasing the number of prizes proportionately as the tournament size increased, subject behavior was similar across replicates. We could not detect a significant difference in the mean effort levels of subjects. This result contradicts the commonly held belief that individuals tend to work less hard in large impersonal organizations than in small ones. It also indicates that tournament replication by managers is not necessarily a suboptimal strategy.

A tournament's prize structure can induce changes in subject effort level. In our six-person tournaments, a change in prize structure from three large prizes (i.e., promotions) and three small, to two large and four small, did not significantly affect the subject effort levels. However, we noted a significant decrease in effort when the tournament's prize structure was changed to four large and two small prizes. This result is striking because if effort decreases in tournaments with many promotions, then the per-unit cost of output significantly increases. A profit-maximizing manager needs to choose a prize structure with the smallest number of promotions that is consistent with the desired output goal.

In our studies, larger organizational size can be used to buffer the firm against discrimination's negative effect on agent output. Consistent with the theory, as we increased the size of our "unfair" or discriminatory laboratory tournaments, mean effort levels increased. This result indicates that the efficiency cost of discrimination is greatest in smaller organizations.

We will proceed as follows: In §2 we review the relevant tournament theory and develop four propositions that theoretically address the above three questions. In §3 we present the experimental design. We present and discuss our results in §4, and offer concluding remarks in §5.

2. Tournaments and Their Equilibria

2.1. A Simple Tournament Model

Consider the following n-person tournament with n identical agents $i = 1, 2, \ldots, n$ each having the same utility functions separable in the payment received and the effort exerted.

$$u(p, e) = w(p) - c(e_i), \quad i = 1, 2, \ldots, n, \tag{2.1}$$

where p denotes the nonnegative payment to the agent, e_i is the agent's nonnegative effort level. The positive and increasing functions $w(\cdot)$ and $c(\cdot)$ are, respectively, concave and convex. Agent i chooses a level of effort from a closed and bounded set on the real line. This effort is not observable to anyone except agent i, but generates an output y_i according to

$$y_i = f(e_i) + \xi_i, \tag{2.2}$$

where the production function $f(\cdot)$ is concave and ξ_i is a random shock drawn independently for each agent i from an identical and continuous density function defined on a common closed and bounded support. All other agents have a similar technology and face an identical decision problem. In this tournament there will be n prizes (i.e., the number of prizes is equal to the number of participants in the tournament) each of which can take one of two values, M or m, with $M > m$. A prize structure $\lambda^T = (\lambda, 1 - \lambda)$ is defined by a fraction, λ, indicating the fraction of large prizes existing in the tournament with $1 - \lambda$ being the fraction of small prizes.

The rules of the tournament are as follows: After outputs are determined for each of the n agents, those λn agents with the largest outputs get the λn large prizes, M, while the remaining $(1 - \lambda)n$ agents get small prizes, m.

In some tournaments, a subset of agents are favored in the sense they do not have to perform as well as disfavored or discriminated against agents in order to win a large prize. This can be modelled by adding a constant k to their output so that whatever their effort, their "effective" output is guaranteed to be larger by k. Such a constant is not added to the output of disfavored agents. O'Keeffe et al. (1984) call such tournaments unfair with k being the discrimination factor, while Lazear and Rosen (1981) call this process, whereby some agents are favored, handicapping.

Given any vector $e = (e_1, \ldots, e_n)$ of effort choices by the agents, agent i's probability of winning M can be denoted by $\pi(e_i, e_{-i})$ where e_{-i} is the vector e with the ith agent's effort level deleted. Thus i's expected payoff from such a choice is

$$\begin{aligned} &Ez_i(e_i, e_{-i}) \\ &\quad = \pi(e_i, e_{-i})w(M) + (1 - \pi(e_i, e_{-i}))w(m) - c(e_i), \\ &\qquad\qquad i = 1, 2, \ldots, n \end{aligned} \tag{2.3}$$

where $\pi(e_i, e_{-i})$ is the probability of winning a large prize for agent i given the effort choices, e_{-i}, of his $n - 1$ competitors.

The above tournament defines a game with payoffs given by (2.3) and a strategy set E given by the feasible set of effort choices, which we assume is a closed interval on the real line. The theory of tournaments restricts itself to the game's pure strategy Nash equilibria; in much of what we do here we will restrict ourselves even further to the symmetric pure strategy equilibria where $e^* = e_1^* = \cdots = e_n^*$. With suitable restrictions on the distribution of random shocks and the utility functions posited above, a unique pure strategy symmetric Nash equilibrium will exist for all the tournaments we examine in our experiments (see Propositions 2 and 4 below). The theory requires the specification of the utility function, the production function, the distribution of shocks, and the prize structure. One specification we use in our experiment is the following:

$$U_i(p_i, e_i) = p - e^2/c, \quad i = 1, 2, \ldots, n. \tag{2.4}$$

$$y_i = e_i + \xi_i, \quad i = 1, 2, \ldots, n, \tag{2.5}$$

where $c > 0$ and ξ_i is distributed uniformly over the interval $[-q, q]$, $q > 0$ and independently across agents. e_i is restricted to lie in $[0, 100]$. In particular, the agents' expected payoff in the tournament is

$$Ez_i(e_i, e_{-i}) = m + \pi(e_i, e_{-i})[M - m] - e_i^2/c \tag{2.6}$$

We call this our *experimental specification* and note it is defined by the parameters $\Gamma = \{M, m, q, c, k, \phi(\xi_i), i = 1, 2, \ldots, n\}$, where M, m, q, k, and c are defined above, and $\phi(\xi_i)$ is a uniform density function determining each independent realization ξ_i.

At the unique interior pure strategy Nash equilibrium, each agent's first-order condition must be fulfilled:

$$\frac{\partial Ez_i}{\partial e_i} = \frac{\partial \pi(e_i, e_{-i})}{\partial e_i} \cdot [M - m] - \frac{2e_i}{c} = 0 \tag{2.7}$$

or

$$\frac{\partial \pi(e_i, e_{-i})}{\partial e_i} \cdot [M - m] = \frac{2e_i}{c}$$

This first-order condition has a simple explanation. On the left-hand side we have the marginal benefit to a tournament participant from increasing his or her effort level. Obviously, this is equal to the increase in the probability of winning caused by the effort increase $\partial \pi(e_i, e_{-i})/\partial e_i$—the marginal probability of winning—multiplied by the net benefit of winning, $[M - m]$. The right-hand side is simply the marginal cost of effort. The second-order conditions guarantee that this is indeed a maximum. Using a uniform distribution with support $[-q, +q]$, at a symmetric equilibrium where $e_i = e_j = e^*$, we know that $\partial\pi(e_i, e_{-i})/\partial e_i = 1/(2q)$ since this follows simply from the fact that we are dealing with a uniform distribution of support $2q$ so that every increase in effort, providing that others do not respond, provides an increase in winning of $1/(2q)$. Hence the first-order condition becomes

$$\frac{1}{2q}\cdot[M-m] = \frac{2e_i}{c} \tag{2.8}$$

which, when solved for the optimal e^*, yields,

$$e^* = \frac{(M-m)c}{4q}. \tag{2.9}$$

Note that whenever the equilibrium marginal probability of winning is $1/(2q)$ for all agents, the equilibrium effort level is defined by (2.9). We will use this fact later. In our experiments we parameterize this model by setting $M = 2.04$, $m = 0.86$, $q = 60$, and $c = 15{,}000$, which determines $e^* = 73.75$.

2.2. Some Theoretical Results on Our Experimental Specification

The above model yields interesting results when we vary the number of tournament participants and the prize structure, and specify the distribution of shocks as uniform. In addition, the introduction of discriminatory behavior by the tournament organizer has interesting implications both for organizational efficiency and social policy. In this subsection we prove a number of simple propositions pertaining to symmetric and asymmetric tournaments. To state our results efficiently we make a few distinctions about types of symmetries and asymmetries with which we are concerned here.

Tournaments are characterized by three factors: the agent characteristics (their utility and cost of effort functions), tournament parameters (the prize distribution λ^T, M, c, and m), and finally the fairness of the rules (whether k is equal to, greater than, or less than zero for any subset of players). We call a tournament *fully symmetric* if all factors mentioned above are symmetric. For example, a tournament is fully symmetric if all players have identical cost functions, $\lambda = 1/2$ so there is an equal number of large and small prizes, and $k = 0$ for all $i = 1, 2, \ldots, n$ so there is no discrimination against (or handicapping of) any group. We are interested in an asymmetry introduced into one of our three factors leaving the other two untouched. We will call a tournament unfair if $\lambda = 1/2$, all agent cost functions are identical, but $k > 0$ for some subset of agents. Likewise, a tournament is *prize-asymmetric* if $k = 0$ for all $i = 1, 2, \ldots, n$, all cost functions for agents are identical, but $\lambda \neq 1/2$, so the number of large prizes is not equal to one-half the number of agents in the tournament. Finally, we call an equilibrium symmetric if $e_i = e^*$ for all $i = 1, 2, \ldots, n$.

With these concepts defined we now state and prove four results which furnish the hypotheses to be tested in our experiments. Note that these propositions are stated for our *experimental specification* with its assumed uniform distribution of shocks. While some generalizations of results are offered which extend them to the family of symmetric unimodal distributions (e.g., normal, etc.), a fully general theory is outside the purview of this paper. Due to size constraints, we do not include the proofs of our propositions. They are available from the authors on request.

Proposition 1. *Let $e^* = (e_1^*, e_2^*, \ldots, e_n^*)$ be a symmetric equilibrium for a fully symmetric tournament with n players using our experimental specification. Then, ceteris paribus, e^* remains a symmetric equilibrium for that tournament as n increases. In addition, holding n constant, e^* remains a symmetric equilibrium for any prize-asymmetric tournament that can be derived from that tournament by allowing λ to vary away from $\lambda = 1/2$ which satisfies the participation constraint of the agents.*

Proof. Available at mansci.pubs.informs.org/ecompanion.html.

Note that Proposition 1 answers both of our earlier questions about organizational replication and prize structure since it shows that, at a symmetric equilibrium, agent effort in tournaments remains constant as the tournament size increases through replication. Or, alternatively, the size is held constant and the distribution of prizes, λ, is varied. (In fact, both n and λn can vary simultaneously with e^* remaining constant.) The second result seems counterintuitive to those who might tend to think in terms of total, not marginal, probabilities. A common mistake is to think that, ceteris paribus, as the number of large prizes in a tournament is increased, equilibrium effort levels should fall since the probability of winning one of those prizes has increased. Hence high effort levels are a costly waste. While it is no mistake to believe that the probability of winning increases when more promotions are available, it is the marginal probability of winning that affects effort, and this probability remains constant. This is the main intuition of the proof.

Proposition 1 is constructed under the assumption that a symmetric interior equilibrium exists. To prove such a result we offer the following proposition.

Proposition 2. *Any fully symmetric or prize asymmetric tournament with our experimental specification has a unique symmetric interior pure strategy equilibrium if $[M-m]c < 4q \cdot 100$ and the participation constraint $\sqrt{c\lambda[M-m]} > e^*$ is satisfied.*

Proof. Available at mansci.pubs.informs.org/ecompanion.html.

Our next result deals with unfair tournaments and the impact that size has on effort levels in such tournaments.

Proposition 3. *As n goes to infinity the equilibrium effort levels in unfair tournaments with $n/2$ disadvantageous agents and $0 \leq k \leq 2q$, asymptotically approaches (from below) the equilibrium effort levels of an otherwise fully symmetric tournament. (Remember, from Proposition 1, e^* is invariant to increases in the size of the tournament.)*

Proof. Available at mansci.pubs.informs.org/ecompanion.html.

Despite its cumbersome nature, the strategy of the proof of Proposition 3 is simple. As we know from the first-order condition (2.8), whenever the marginal probability of winning in equilibrium is equal to $1/(2q)$ for all agents, the equilibrium effort level for agents in a symmetric equilibrium is $e^* = (M - m)c/(4q)$. Hence, if we can show that, as n gets large, the equilibrium marginal probability of winning in an unfair tournament increases monotonically from below toward $1/(2q)$ (and converges to $1/(2q)$ for all agents as $n \to \infty$), we would have proven our result. However, as is demonstrated in the proof of Proposition 3, the equilibrium marginal probability of winning in an unfair tournament with $n/2$ large prizes and $n/2$ favored agents (receiving a $k > 0$) can be written as $\partial(\Pr(i \text{ wins}))/\partial e_i = 1/(2q) - k^{n/2}/(2q)^{n/2+1}$ for all agents. With $k < 2q$, this term increases monotonically as n increases, and converges to $1/(2q)$, as $n \to \infty$.

This proposition answers the third question posed in §1 since it proves that as n becomes larger, the efficiency loss of discrimination gets smaller, since equilibrium effort levels increase to their nondiscriminatory levels.

Proposition 4. *Assuming the appropriate second-order conditions are satisfied, all unfair tournaments satisfying our experimental specification have a unique symmetric pure strategy equilibrium.*

Proof. Available at mansci.pubs.informs.org/ecompanion.html.

2.3. Hypotheses To Be Tested

Propositions 1 and 3 define a set of three null hypotheses which serve as the basis for the discussion of experimental results. In testing each hypothesis we compare our data to the predictions of the symmetric pure strategy equilibrium. While mixed and perhaps asymmetric equilibria may also exist, we will see that our symmetric equilibrium results do a good job of organizing the data. Hence we pay no attention to these other equilibria if, in fact, they even exist.

We state our three hypotheses in the order by which our experimental design tests them.

Hypothesis 1 examines the effects of tournament replication. Tournament theory predicts that behavior is invariant to any degree of replication. That is, agent effort is homogenous of degree zero, for proportionate increases in tournament size. Letting $e^*_{\text{Exp}\,j}$ denote the symmetric equilibrium effort level of subjects in tournament j we have:

Hypothesis 1. $e^*_{\text{Exp}\,1} = e^*_{\text{Exp}\,2} = e^*_{\text{Exp}\,3}$.

Hypothesis 2 examines whether behavior is invariant to prize structure when size is held constant. We test this by changing the prize structure in three 6-person tournaments, and observe how changes in prize structure affect subject effort. The theory predicts identical behavior across structures.

Hypothesis 2. $e^*_{\text{Exp}\,3} = e^*_{\text{Exp}\,4} = e^*_{\text{Exp}\,5}$.

Hypothesis 3 looks at the effect of tournament size on subject effort in unfair tournament. The theory predicts that when discrimination is present, effort is greater in larger tournaments. In Experiments 6 to 8 we hold the level of discrimination constant, while we increase organizational size.

Hypothesis 3. $e^*_{\text{Exp}\,8} > e^*_{\text{Exp}\,7} > e^*_{\text{Exp}\,6}$.

These hypotheses clearly have managerial implications in the efficient use of internal resources. If Hypotheses 1 and 2 are confirmed, they would offer support for tournaments as flexible incentive structures. For example, managers should expect the same effort from agents if they were organized into one large 100-person tournament with 50 promotions, or ten 10-person tournaments with 5 promotions. So, if true, Hypothesis 1 predicts that for tournaments approaching fully symmetric, agent output should not vary as organizational size increases. Clearly, the choice of tournament size should then depend on which type of organization has lower administrative costs.

Hypothesis 2 predicts that effort is invariant to changes in the prize structure. Subject to retaining the interior equilibrium of the tournament, managerial choice of prize structure should not affect agent effort. Since total labor cost is defined as the sum of tournament prizes, one managerial implication is that a profit-maximizing manager wants to design a tournament with as few large prizes (i.e., promotions) as possible.

Hypothesis 3 predicts that the negative effects of discrimination on agent output can be neutralized by using larger tournaments. Managers who suspect any favoritism, therefore, should construct larger tournaments. They do so because in unfair tournaments agents increase their effort levels as tournament size increases. Hypotheses 2 and 3 suggest that the incentive structure of profit-maximizing organizations should resemble tournaments with a large number of agents competing for relatively few promotions.

3. The Experiments Performed and the Experimental Design

We designed eight experiments to examine the effects of organizational size and tournament design. Experimental parameters are shown in Table 1. We held relevant tournament parameters constant across all eight experiments. These include the decision number and random number ranges, the cost-of-effort function, and the fixed payment differential between high and low prizes (M and m).

Experiments 1 to 3 used fully symmetric tournaments to test Hypothesis 1. All experiment parameters across these three experiments are constant except for tournament size. Size increased from 2-person tournaments in Experiment 1 to 4-person in Experiment 2, and 6-person in Experiment 3.

Experiments 3 to 5 examined the predictions of Hypothesis 2. These 6-person experiments were identical, except for changes in their prize structure. Experiment 3 used three large prizes, Experiment 4 used four large prizes, and Experiment 5 used two large prizes in their respective price structures.

We tested Hypothesis 3 in Experiments 6 to 8. All experiments were conducted with identical unfair, asymmetric tournaments. As with Experiments 1 to 3, only the tournament size differed. Experiment 6 used two-person, Experiment 7 used four-person, and Experiment 8 used six-person tournaments. All had $\lambda = 1/2$.

3.1. Experimental Procedures

Two hundred and eighty six subjects were recruited from the New York University and University of Pennsylvania undergraduate classes. They were randomly assigned to time slots, and instructed to meet at the behavioral laboratory on their respective campuses. The laboratory consisted of cubicles that made it difficult to see others during the experiment. Subjects were randomly assigned to cubicles and instructions were publicly read to them (See online appendix). The instructions basically said the following: For this experiment, subjects would be randomly assigned with a specified number of other subjects (one, three, or five, depending on the tournament size). These subjects would be their "group members." Group members remained the same during the entire experiment, and their physical identities were not revealed. Subjects were told the amount of money they earned was a function of their decisions, their group members' decisions, and the realization of a random variable as described by the rules of the tournament. They were then given cost-of-effort tables and told that all subjects had identical tables and instructions. In each round, their task was to chose an effort level (to which a random component was added by the computer or by pulling a chip from a bag of chips, if the experiment was a two-person experiment and hence performed by hand). After each round, they were shown their effort, their random realization, and whether they earned a large or a small fixed payment. They learned nothing about the effort levels or the random realizations of their cohorts. All parameters were common knowledge except the identity of pair members.

All two-person tournaments were conducted manually, while four- and six-person tournaments were run using computer terminals. A more detailed set of experimental procedures for our two-person experiments can be found in Bull et al. (1987).

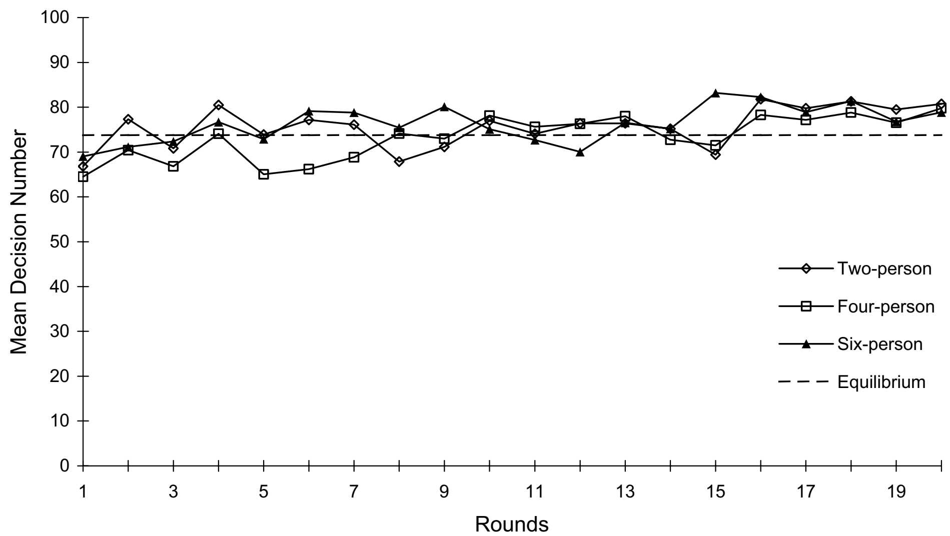
4. Results

To organize our data we present experimental results to test our three hypotheses. Experimental results are shown in Figures 1 to 4, and Tables 2 to 5. Figures 1, 2, and 4 show the round-by-round subject mean effort

Table 1 Experimental Parameters

Experiment	Tournament size	Decision number range	Cost function	Random range	M	m	Number of M	Predicted equilibrium advan.—disadvan.	Number of subjects
Symmetric tournament									
1	2	(0–100)	$e^2/15{,}000$	$(-60, 60)$	\$2.04	\$0.86	1	73.75	24
2	4	(0–100)	$e^2/15{,}000$	$(-60, 60)$	\$2.04	\$0.86	2	73.75	52
3	6	(0–100)	$e^2/15{,}000$	$(-60, 60)$	\$2.04	\$0.86	3	73.75	24
4	6	(0–100)	$e^2/15{,}000$	$(-60, 60)$	\$2.04	\$0.86	4	73.75	24
5	6	(0–100)	$e^2/15{,}000$	$(-60, 60)$	\$2.04	\$0.86	2	73.75	24
Asymmetric tournaments—1/2 of subjects advantaged ($K = 25$)									
6	2	(0–100)	$e^2/15{,}000$	$(-60, 60)$	\$2.04	\$0.86	1	58.39 58.39	18
7	4	(0–100)	$e^2/15{,}000$	$(-60, 60)$	\$2.04	\$0.86	2	70.55 70.55	36
8	6	(0–100)	$e^2/15{,}000$	$(-60, 60)$	\$2.04	\$0.86	3	73.08 73.08	42

Figure 1 **Mean Effort Choices in Two-, Four-, and Six-Person Fully Symmetric Tournaments**



choices across experiments. Figure 3 shows the organizational cost per unit of effort in six-person symmetric tournaments. Table 2 presents summary statistics of observed behavior in Experiments 1 to 3. Tables 3 and 4 do the same for Experiments 3 to 5. Table 5 shows summary statistics of observed behavior in Experiments 6 to 8.

4.1. Hypothesis 1: Organizational Replication of Symmetric Tournaments

Hypothesis 1 states that observed behavior in fully symmetric tournaments should be invariant to tournament size. Hence, we expect observed effort levels in Experiments 1, 2, and 3 to be identical ($e^*_{\mathrm{Exp}\,1} = e^*_{\mathrm{Exp}\,2} = e^*_{\mathrm{Exp}\,3} = 73.75$). Figure 1 and Table 2 compare subject behavior across 3 organizational replications—two-person, one large prize; four-person, two large prizes; and six-person, three large prizes. Figure 1 clearly shows very similar behavior across the replications. To test Hypothesis 1 we investigate whether observed behavior deviates from predicted.

First, we test whether mean subject choices in each tournament significantly differ from that predicted (73.75). A round-by-round Wilcoxon signed rank test does not reject the hypothesis that observed effort levels in each replicate came from a population with a mean of 73.75.[1] As shown in Table 2, the highest observed deviation from predicted effort in rounds 1 to 10 is 9.44. In all replicates, the worst deviation from predicted occurs in round 1. The learning patterns are similar: by rounds 11 to 20, the mean observed deviation is only 4.6.

We then test whether behavior is similar across replicates. We find no significant difference in behavior across any decision round. A pairwise, round-by-round Mann-Whitney confirms no significant difference; we cannot reject the hypothesis of effort drawn from identical populations. These data strongly suggest behavior is constant across replicates.[2]

This finding supports Hypothesis 1. Subjects behaved similarly across prize structure replication. As long as the percentage promoted remains constant with organizational growth, behavior does not change. Size does not matter, if promotion rates remain constant.

Though participants may be indifferent to tournament size, administrators should care. The similar behavior in our symmetric replicates implies a constant cost per output unit. The resource efficiencies of the replicates are virtually identical. Since tournament size and prize structure are directly controlled by upper-level managers, there is a definite link between tournament design and resource efficiency. Clearly, if the fixed costs of running tournaments are constant across size, then managers should prefer larger to smaller tournaments, to reduce the costs. This is especially true if economies of scale exist. We also find another link between efficiency and size below (Hypothesis 3). When favoritism is present, larger size can buffer the negative effects on effort.

[1] The Wilcoxon signed-rank test requires a symmetric distribution of data. Using a Kolmogorov-Smirnov test, we could not reject the hypothesis that the data were drawn from a normal, hence symmetric, distribution. All statistical tests used a significance level of 5%.

[2] Our results are based on small samples. We ran power tests to obtain an indication of strength of statistical inference. Power values are 0.31 for Experiment 1, 0.93 for Experiment 2, and 0.75 for Experiment 3. We also pooled results across the three experiments. Using similar tests as reported above, we could not reject the hypothesis that effort was drawn from a distribution with a mean of 73.75 in 19 of 20 rounds.

Figure 2 **Mean Effort Choices in Six-Person Prize-Asymmetric Tournaments** ($\lambda = 1/2, 1/3, 2/3$)

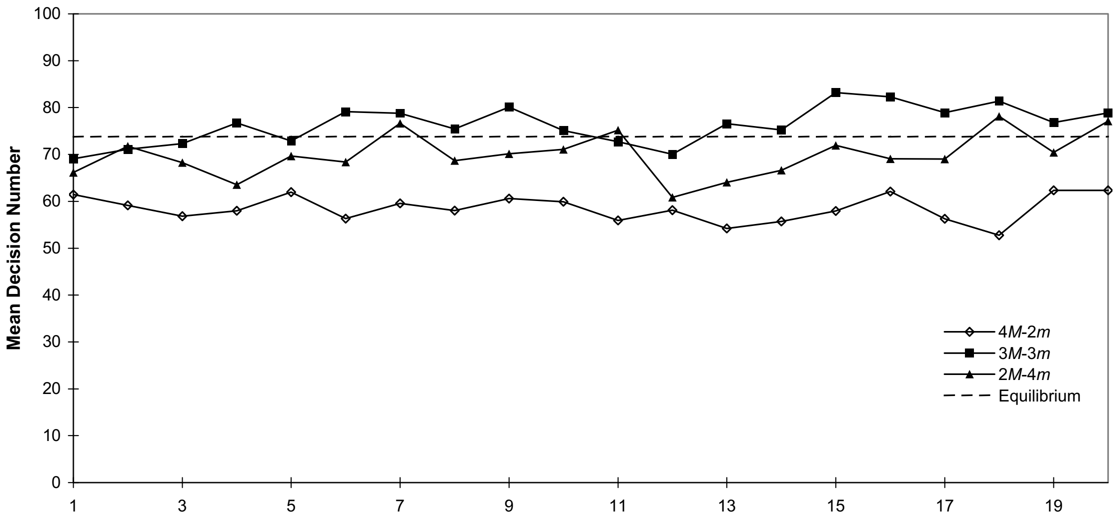
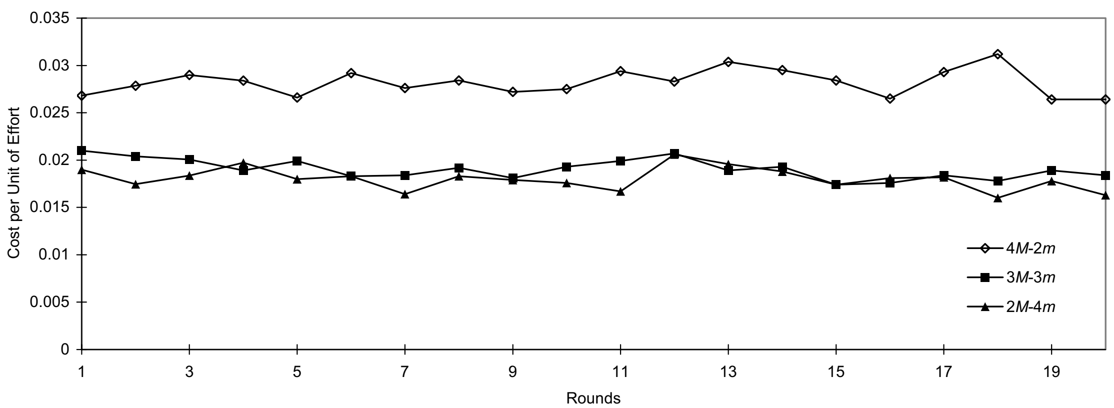


Figure 3 **Organizational Cost per Unit of Effort in Six-Person Prize-Asymmetric Tournaments**



4.2. Hypothesis 2: Behavior and Prize Structure

Hypothesis 2 predicts similar identical behavior when size is held constant, and prize structure is changed. So we should observe similar effort levels in Experiments 3 to 5 (6-person tournaments with different prize distributions of $3M$-$3m$, $2M$-$4m$, and $4M$-$2m$). Figure 2 and Table 3 show observed behavior. We test for similarity by applying a round-by-round Mann-Whitney test. In Experiments 3 ($3M$-$3m$) and 5 ($2M$-$4m$), we cannot reject the hypothesis of efforts drawn from identical populations. This suggests similar behavior across these structures. However, this result does not extend to behavior in Experiment 4 ($4M$-$2m$). As shown in Figure 2, effort levels in the $4M$-$2m$ prize structure are significantly lower than in the other two. The Mann-Whitney test confirms this behavior in every round.

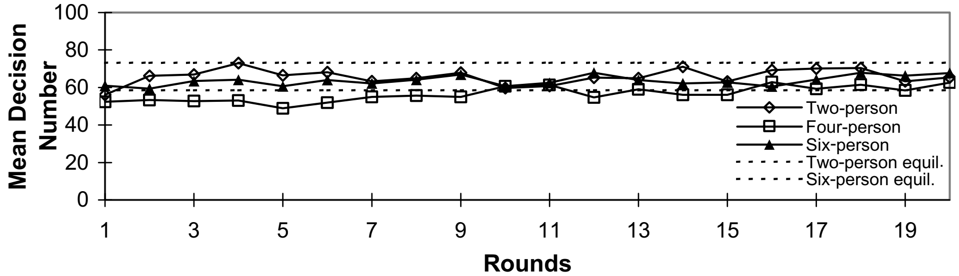
We next test whether observed behavior is as predicted. The predicted effort in each of these tournaments is 73.75. As shown in §4.1, behavior in Experiment 3 ($3M$-$3m$) is similar to that predicted in every round. A similar test reached the same conclusion for Experiment 5 ($2M$-$4m$) results. Behavior in Experiment 4 ($4M$-$2m$) is not as predicted. Effort levels in every round are significantly lower than predicted.[3]

In our setting, tournaments with a relatively high proportion of big prizes elicit low effort. Hence, they are also more costly. This suggests that when promotions are plenty, agents will shirk.

4.2.1. Organizational Costs. Managers interested in efficient resource use want to elicit similar behavior at lower costs. One measure of managerial efficiency is the organizational cost per effort unit. We note that while the prize structures in Experiments 3 to 5 are predicted to produce the same effort level (73.75), their costs differ because of prize structure. Theoretically, the predicted organizational cost per effort unit in Experiment 3 is \$0.0197, for Experiment 4 it is \$0.0223, and for Experiment 5 it is \$0.0170.

[3] Power tests revealed a value of 0.75 for Experiment 3, 0.27 for Experiment 4, and 0.75 for Experiment 5.

Figure 4 Mean Effort Choices in Two-, Four-, and Six-Person Asymmetric Tournaments



So in both prize asymmetric tournaments (4*M*-2*m*, 2*M*-4*m*) we see similar effects for opposite reasons. In the 4*M*-2*m* tournaments subjects may believe it is easier to win, so they reduce effort levels; in the 2*M*-4*m* design they think it is more difficult to win, so they slightly reduce effort levels.

If observed behavior (and hence organizational costs) differ from that predicted at the equilibrium, an interesting empirical question is, How do they differ? We present observed organizational costs in Figure 3 and Table 4. Obviously, from looking at Figure 3 we see that the 4*M*-2*m* design leads to the highest cost per unit of effort, both because of the higher number of large fixed payments and because of the lower effort levels exerted (see Figure 2). Further, as shown in Table 4, in all of the last 10 rounds of Experiment 4, the observed cost per effort unit is higher than predicted in the 4*M*-2*m* design.

It is interesting to compare actual costs between the 3*M*-3*m* and 2*M*-4*m* design. We previously discussed how subjects exerted slightly more effort in the 3*M*-3*m* relative to the 2*M*-4*m* design. This would seem to imply that managers are better off using a symmetric (equal number of high and low payments) compensation design. However, there is obviously an inherent advantage in offering a smaller number of high payments, since this design reduces the tournament's cost. We see the predicted difference between the two designs is 0.0027 per unit of effort. Table 4 shows that, although the actual cost difference is smaller than that predicted (since the 3*M*-3*m* design elicits slightly higher effort), the 2*M*-4*m* design still results in lower costs. That is, while the cost differential between the two designs is reduced, the 2*M*-4*m* design still produces a lower cost per unit of effort—an average of 0.0179 over the last 10 rounds compared with 0.0187 for the 3*M*-3*m* design. However, over the last 10 rounds, we could find no statistical difference in organizational cost across the two designs.

4.3. Hypothesis 3: Discrimination and Tournament Size

We test the effect of size in unfair tournaments by fixing the degree of discrimination ($k = 25$), and increasing size from two-person to four-person, and finally to six-person. Hypothesis 3 states that effort should monotonically increase in size. Hence, the loss in organizational efficiency due to discrimination decreases in tournament size. The predicted effort level in Experiment 6 is 58.39, in Experiment 7 it is 70.55, and in Experiment 9 it is 73.08. There should be no difference in behavior across advantaged and disadvantaged subjects. Figure 4 and Table 5 compare subject behavior across these experiments.

Results suggest that behavior in the four- and six-person tournaments is similar to that predicted. In both Experiments 7 and 8 a round-by-round Mann-Whitney test for rounds 11 to 20 shows the following: There is no significant difference in observed effort level across advantaged and disadvantaged subjects. We cannot reject the hypothesis that effort was drawn from a population with a mean of 70.55 in Experiment 7 and a mean of 73.08 in Experiment 8. Finally, though the predicted means are within 3.5% of each other, we can detect some significant difference in effort choices across the two experiments. Mean effort choice in the four-person tournament is greater than that of the six-person tournament in only 1 round. It is significantly lower in 5 of the last 10 rounds.

Table 2 Organizational Replications of Symmetric Tournaments

		Mean decision numbers			Max. deviation		Mean deviation	Mean standard deviations		
Experiment	Predicted	Rounds 1–10	Rounds 11–20	Round 20	Rounds 1–10	Rounds 11–20	Rounds 11–20	Rounds 1–10	Rounds 11–20	Round 20
Two-person $N = 24$	73.75	73.87	77.91	80.75	6.95	9.0	5.3	25.04	24.75	23.46
Four-person $N = 52$	73.75	70.12	76.47	79.77	9.25	6.0	3.7	9.83	9.49	5.51
Six-person $N = 24$	73.75	75.07	77.59	78.85	9.44	9.4	4.8	7.04	7.88	2.8

Table 3 Six-Person Symmetric Tournaments—Different Compensation Designs

Experiment		Predicted	Mean decision numbers			Maximum deviation		Mean deviation	Mean standard deviations		
			Rounds 1–10	Rounds 11–20	Round 20	Rounds 1–10	Rounds 11–20	Rounds 11–20	Rounds 1–10	Rounds 11–20	Round 20
$2M$-$4m$	$N = 18$	73.75	69.42	70.25	77.10	10.22	12.88	5.3	6.18	7.84	15.42
$3M$-$3m$	$N = 24$	73.75	75.07	77.59	78.85	6.95	9.44	4.8	7.04	7.88	2.83
$4M$-$2m$	$N = 24$	73.75	59.18	57.77	62.33	17.42	20.97	16.1	15.7	19.63	16.18

However, behavior in the two-person unfair tournament is significantly different from that predicted. We find that effort levels across the advantaged and disadvantaged significantly differ in 7 of the last 10 rounds. Mean effort levels in rounds 11 to 20 are significantly higher than the predicted 58.39. In fact, Figure 4 shows that for rounds 11 to 20 mean effort level is always higher in the two-person relative to the four-person tournament. In comparing mean effort levels in the two-person relative to the six-person tournament, we see they are higher more than half the time. Round-by-round Mann-Whitney tests reveal no significant differences in behavior.[4]

In summary, our results are partially counter to the theoretical predictions. While theory predicts an increase in effort levels as the size of the organization increases, our results support this prediction only when going from a four- to a six-person tournament; they do not support it when moving from a two- to four-person tournament or even from a two- to six-person tournament; effort levels in a two-person tournament were significantly higher than predicted.

5. Discussion

There is growing evidence that firms use promotion tournaments to incentivize agents, yet little research has examined behavior within tournament settings. We do so with a series of laboratory tournaments, whose design allows us to examine how subject output is affected by tournament design. In firms, these design parameters are managerially controlled. Since in tournaments output and prize structure define organizational cost, design is one way managers control the use of internal resources.

Our findings are generally supportive of tournament theory. In many of our sessions, we observe behavior resembling that predicted by the theory. For a few predictions, we find little support. We discuss these findings in greater detail below, and note the managerial implications. In discussing these implications, we recognize the issues that are raised when one tries to extrapolate laboratory findings to the business world. So the implications we draw are speculative, and conditional on our findings extending to larger and more complex tournaments.

Table 4 Observed Cost per Unit of Effort

Compensation design	Predicted cost/unit	Average Rounds 1–10	Average Rounds 11–20	Round 20 cost	Number of Rounds below predicted in rounds 11–20
$2M$-$4m$	0.0170	0.0181	0.0179	0.0163	3
$3M$-$3m$	0.0197	0.0194	0.0187	0.0184	8
$4M$-$2m$	0.0223	0.0279	0.0286	0.0264	0

5.1. Design Issues and Output

Our findings indicate that tournament design does impact output and, hence, organizational cost. We find support that tournament replication in response to organizational growth is not per se suboptimal. In fact, if the tournament design is optimal for the given organization size, then a replication strategy can be an optimal strategy for managers. In our tournaments, subject behavior is invariant to tournament size, when the proportion of high prizes is kept constant.

Prize structure also affects subject behavior. Behavior in the $3M$-$3m$ and $2M$-$4m$ tournaments was as predicted. We found no significant difference in subject behavior across the two experiments. However, subjects in the $4M$-$2m$ experiment significantly decreased their effort. Observed effort levels chosen by subjects were significantly lower than the predicted 73.75. We find this behavior interesting. It seems when there is a high percentage of large prizes, subjects tend to shirk. We speculate one reason for this behavior is that some subjects may not sufficiently recognize the difference between the marginal and total probabilities of winning.

There is an incentive in tournaments to reduce effort if the probability of winning increases. While the $4M$-$2m$ design has the same marginal probability of winning as the $2M$-$4m$ and $3M$-$3m$ designs (see Proposition 1), of the three designs, $4M$-$2m$ has the highest probability of winning (0.66). To illustrate this point, imagine you are an assistant professor competing with four others of identical ability for tenure. Your department chair tells each candidate that four of you will receive tenure. That is, the tenure rate is

[4] The power associated with these experiments were 0.37 for Experiment 6, 0.51 in Experiment 7, and 0.73 in Experiment 8.

Table 5 Organizational Replications of Asymmetric Tournaments

		Mean decision numbers			Maximum deviation		Mean deviation	Mean standard deviations		
Experiment	Predicted	Rounds 1–10	Rounds 11–20	Round 20	Rounds 1–10	Rounds 11–20	Rounds 11–20	Rounds 1–10	Rounds 11–20	Round 20
Two-person $N = 18$	53.39	65.27	66.58	65.33	14.66	12.61	7.95	25.16	28.91	32.22
Four-person $N = 36$	70.55	53.89	59.19	62.58	21.61	15.8	11.35	22.14	23.62	24.50
Six-person $N = 42$	73.08	62.65	64.60	67.76	13.69	12.56	8.48	13.21	14.91	12.49

equal to 80%. With such odds, a boundedly rational decision maker may surmise that he could slightly reduce effort. After all, one may think that while one has to work very hard to be the first or second best, one doesn't have to work that hard to be fourth best (out of five). A slight lowering of effort would reduce the probability of winning (from its current 80%), but it would also increase the payoff, conditional on winning, since reducing effort decreases the candidate's cost of effort.

So if subjects are anchoring on the total probability of winning in the $4M$-$2m$ design, they believe the probability of winning is 66.6%. They may decide to lower their effort level to decrease the associated costs. For example, a drop in effort from 73.75 to 65 results in a 22% decrease in cost of effort. The subjects probably understand they are lowering their probability of winning, but even if they lose, the lower effort results in a higher payoff. Clearly, subjects are willing to trade off lower probabilities of winning (reducing them toward 50%), for reductions in associated effort costs.

These results have managerial implications for the proper design of organizational incentives. One crucial task that managers face in allocating internal resources is the construction of effective incentive structures. Such structures need to modify agent behavior in the desired way, and minimize the associated labor costs. Our theoretical results indicate that, when random shocks are from a symmetric unimodal distribution, then equilibrium effort levels are invariant to the prize structure (for interior equilibria). Empirically, we also find that when too many large prizes are dangled before agents, those agents have a tendency to reduce their effort levels. Hence, the optimal prize structure is one that minimizes the number of promotions, conditional with maintaining the individual participation constraint. So a steeper organizational pyramid seems better for agent motivation, as long as the number of possible promotions doesn't become so low that agents drop out of the tournament and provide no effort.

5.2. Discrimination and Tournaments

We find evidence that tournament size can reduce the inefficiency effects of discrimination, thereby reducing its opportunity costs. This finding suggests that if efficiency is the sole reason why society is concerned about discrimination in the workplace, then governmental enforcement of discrimination laws are best aimed at small, rather than large firms, since the costs are greater in the smaller firms. However, as Becker (1957) has suggested, if employers have a taste for discrimination which is not exercised if the efficiency costs are high, then we might expect to see more discrimination in larger firms since the associated opportunity costs of discrimination are smaller in them. This would imply that the government should monitor large firms more closely since the market alone can police smaller firms whose managers have a costly taste for discrimination.

Evidence appears to support the idea that the government monitors larger firms with their lower opportunity costs of discrimination. For example, firms with fewer than 25 employees are exempt from filing Equal Employment Opportunity Commission (EEOC) reports. Medoff (1985) provides evidence that this emphasis on large firm enforcement has been effective. He reports that the rate of discrimination has significantly decreased since equal opportunity policies were enacted in the 1960s.

5.3. Future Research

These experiments open the door to a much more ambitious future program of research. For example, it would be interesting to investigate the impact of different disturbance distributions on the performance of agents in these tournaments. Besides changing the prize distribution and investigating its efficiency consequences, we can also investigate what happens when the fraction of agents in a tournament who are discriminated against changes. More precisely, in the experiments above when discrimination exists one-half of the population of subjects are discriminated against. Hence the fraction of advantaged and disadvantaged subjects is equal. What would happen, however, if we let the fraction of disadvantaged subjects increase or decrease away from one-half? Would being one among many disadvantaged subjects increase or decrease one's effort in a tournament away from what it would be in that same tournament if the advantaged and disadvantaged subjects were equal in number? Many other such questions also exist.

An online appendix to this paper is available at mansci.pubs.informs.org/ecompanion.html.

Acknowledgments
The authors thank the C. V. Starr Center for Applied Economics at New York University and the Reginald Jones Center for Management Policy, Strategy, and Organization at the University of Pennsylvania for their financial support. They thank two anonymous referees and the associate editor as well as Colin Camerer, Yaw Nyarko, Tech Ho, and Steve Hoch for their advice and suggestions. All errors are those of the authors alone.

References

Baker, G., M. Gibbs, B. Holmstrom. 1993. Hierarchies and compensation: A case study. *Eur. Econom. Rev.* **37** 366–378.

Baker, G., M. Gibbs, B. Holmstrom. 1994. The internal economics of the firm: Evidence from personnel data. *Quart. J. Econom.* **109** 881–919.

Becker, G. S. 1957. *The Economics of Discrimination*, 2nd. ed. University of Chicago Press, Chicago, IL.

Bognanno, M. 2001. Corporate tournaments. *J. Labor Econom.* **19** 290–315.

Bull, C., A. Schotter, K. Weigelt. 1987. Tournaments and piece rates: An experimental study. *J. Political Econom.* **95** 1–33.

Conyon, M., S. Peck, G. Sadler. 2001. Corporate tournaments and executive compensation: Evidence from the U.K. *Strategic Management J.* **22** 805–815.

Lambert, R., D. Larcker, K. Weigelt. 1993. The structure of organizational incentives. *Admin. Sci. Quart.* **38** 438–461.

Lazear, E., S. Rosen. 1981. Rank-order tournaments as optimum labor contracts. *J. Political Econom.* **89** 841–864.

Leonard, J. 1990. Executive pay and firm performance. *Indust. Labor Relations Rev.* **43** s13–s29.

Medoff, M. 1985. Discrimination and the occupational progress of Blacks since 1950. *Amer. J. Econom. Sociology* **44** 295–303.

O'Keefe, M., K. Viscusi, R. Zeckhauser. 1984. Economic contests: Comparative reward schemes. *J. Labor Econom.* **2** 27–56.

Reiss, R. D. 1989. *Approximate Distributions of Order Statistics*. Springer-Verlag, Berlin, Germany.