Mixed Tournaments, Common Shocks, and Disincentives: An Experimental Study

Wu, Steven; Roe, Brian and Sporleder, Thomas

September 2006

Mixed Tournaments, Common Shocks, and Disincentives: An Experimental Study

Steven Y. Wu, Brian E. Roe, and Thomas Sporleder*

Abstract

Experimental economics is used to investigate two important hypotheses proposed in the economics literature on tournaments. Specifically, we test for a hypothesized "disincentives effect" which can occur in tournaments with mixed ability agents. We also test the well known hypothesis that, when common shocks are an important source of risk, tournaments can filter out this common shock and reduce earnings risk to workers. We find that disincentive effects arose in our tournament experiments, although these effects are not as strong as we predicted in our theoretical model and simulations. We also find that tournaments can be very effective at reducing earnings variation when common shocks are important. Taken together, these results suggest that the benefits of risk reduction from eliminating common shocks might outweigh the disincentive effects arising from mixed tournaments. We also find that the *difference* in average earnings between low and high ability agents is greater under tournaments than under absolute performance contracts.

Keywords: mixed tournaments, incentives, relative performance contracts, experimental economics.

JEL Classification: C91 D01, D81, D82, D86

---

* The authors are assistant professor, associate professor and professor, respectively, at The Ohio State University. Wu is contact author and can be reached at wu.412@osu.edu. The authors gratefully acknowledge funding from the USDA/NRICGP program for Markets and Trade or Rural Development, Award no. 2003-35400-12887 and the Ohio Agricultural Research and Development Center. The authors would like to thank Aaron Stockberger, Tom Keehner, and Chris Dunn for excellent research assistance. We would also like to thank John Kagel for valuable suggestions that helped improve our experimental design.

Mixed Tournaments, Common Shocks, and Disincentives: An Experimental Study

Economists have devoted considerable attention to the study of alternative incentive schemes designed to motivate workers and align their interests with those of the firm. In particular, tournament incentive schemes have received considerable attention in the literature because they are pervasive in the real world (Bull, Schotter and Weigelt 1987). Under tournaments, workers are ranked against other workers and the level of their total pay depends on their relative ranking. Tournaments are often used to determine compensation for salespeople, executives, and for farmers raising broilers and tomatoes. Moreover, political elections, promotions within firms, grading on curves in schools, and competition for positions in firms or sports teams can be interpreted as tournaments.

Two important and well known predictions emerge from the economics literature on tournaments. First, when there are common shocks that affect the performance of an entire cohort of workers/agents, then tournaments can be highly effective incentive devices. Under tournaments, workers are ranked against other workers and the level of their total pay depends only on their relative ranking so that the impact of the common shock on pay is essentially filtered out. Green and Stokey (1984) show theoretically that, in the absence of common shocks, absolute performance contracts Pareto-dominate tournaments but the reverse might be true when common shocks are important.[1] Second, there can be some disincentive effects from using tournaments when agents are of mixed ability (Lazear and Rosen 1981; Knoeber and Thurman 1994). Intuitively, when high ability agents compete against low ability agents, high ability agents can shirk and still win with high probability whereas low ability agents have a low

[1] Under absolute performance contracts, a worker's performance is measured against some fixed standard(s) and total pay depends on whether performance meets or exceeds standard(s). One can construct an absolute performance contract where attempting to meet or exceed standard(s) has similar incentive effects as trying to improve one's rank under tournaments, although one's ranking under tournaments also depends on how co-workers perform.

probability of winning even if they work hard. Thus, neither agent has an incentive to work hard in equilibrium.

Ironically, most experimental studies on tournaments (e.g. Bull, Schotter and Weigelt 1987; Schotter and Weigelt 1992; Nalbantian and Schotter 1997; van Dijk, Sonnemans, and van Winden 2001) ignore common shocks which prevents inferences to be made about the ability of tournaments to reduce payment risks when common shocks are present. Previous experimental studies do, however, examine tournaments with agents of mixed abilities (Bull, Schotter and Weigelt 1987; Schotter and Weigelt 1992; van Dijk, Sonnemans, and van Winden 2001). For the most part, low ability subjects in these studies tend to work as hard or harder than predicted under tournament theory. This led van Dijk, Sonnemans, and van Winden (p. 209) to suggest that "…theories which do not take into account factors such as a desire to outperform others irrespective of monetary gain, do not fare very well," which implies that the disincentive effects of tournaments may be offset by other extrinsic psychological factors. However, it is difficult to draw definite conclusions from these earlier studies as they were not designed to explicitly test for disincentive effects.

In this study, we attempt to fill the research gaps left by earlier experimental studies by testing for the disincentive effect of mixed tournaments and the risk reducing properties of tournaments in the presence of common shocks. It is important to empirically examine these aspects of tournaments because there is a possible tension between the positive effects of tournaments (reduce payment risk) and the disincentive effects from mixing agents in tournaments. To understand this tension, note that tournaments are desirable theoretically because they can filter out the common shock which would reduce workers' risk exposure. A straightforward application of Holmstrom's (1979) informativeness principle then suggests that the firm can provide stronger incentives under tournaments relative to absolute performance contracts, thereby enabling firms to either implement higher effort or lower the cost of implementing a particular effort profile.[2] This advantage of tournaments over absolute performance contracts increases with the importance of the common shock. However, the wrench in the cog is that reducing risk means that high ability workers become more likely to win the tournament than low ability workers. This is because when the relationship between performance and effort is noisy, low ability agents can still get lucky and win so low ability agents still have a reasonable probability of finishing ahead of the high ability agent. However, as the common shock becomes more important (i.e. common shock variance increases), the tournament eliminates a higher proportion of total risk thereby significantly reducing the luck component. Thus, when common shocks are important, the low ability agent is less likely to win the tournament, which increases the disincentive effect. If this disincentive effect is strong enough, it is possible for absolute performance contracts to be preferred to tournaments *in environments where the common shock is dominant*!

Lazear and Rosen suggest some mechanisms for mitigating the disincentive effects of mixed tournaments. One possibility is that tournament organizers can screen potential employees to create a more homogeneous pool. However, in practice, mixed pools may be unavoidable since there is typically a shortage of high ability workers to fill all positions so firms may be forced hire some lower ability workers to ensure that personnel needs are met.[3] Lazear

---

[2] In an interesting twist, Hvide (2002) showed that workers may prefer more risk because it allows them to lower effort. Thus, when their effort can influence the distribution of pay, then they take too much risk and exert too little effort.

[3] For example, teams in Major League Baseball have adopted an extremely intensive sorting system to screen players. A player who has reached the major leagues was most likely heavily evaluated by professional scouts since he was in high school. From there, he may play college baseball followed by minor league baseball. Only if the player survives minor league baseball, where competition is fierce, does the player ever reach the major leagues. Yet, despite this intense sorting process, there is substantial heterogeneity in major league talent. One possible

and Rosen also suggest that handicapping tournaments to increase the competitiveness of less able players can also mitigate the disincentive effects of mixed tournaments. However, handicapping is rarely feasible in most work places as they may be perceived to be unfair and high ability workers may threaten to quite. Handicapping may, however, be feasible in some special settings, such as when firms are trying to fill desirable positions by either promoting from within or hiring externally (e.g. Chan 1996; Tsoulouhas, Knoeber, and Agrawal, forthcoming). Consequently, in most cases, complete and perfect sorting and the use of handicap schemes may not be feasible. Thus, it is important to study the benefits and costs of using tournaments when agents/workers are of heterogeneous ability.

An important aspect of our study is that we show through our modeling simulations that, as the common shock becomes more important (its variance increases), the probability of the high ability agent winning the tournament increases while the probability of the low ability agent decreases. In the treatment where the common shock variance is highest, the low ability agent has such a low probability of winning that he/she is better off "quitting" by choosing zero effort. Thus, in contrast to previous studies, we designed an experimental treatment that allows us to explicitly test for the disincentive effects of mixed tournaments. Moreover, by comparing payment variation across experiments with different common shock variances, we can assess whether tournaments reduce payment risk as predicted. If tournaments induce both disincentives and a reduction in payment risk, then contract designers must carefully balance these two effects to obtain optimal tournament contracts in mixed ability environments. Essentially, keeping agents exposed to a sufficient amount of payment risk would mitigate the disincentive effect.

explanation for this is that there is only a handful of elite players for each position so that many teams must make due with the "best that they've got" from their minor league system.

Our results suggest that, in the treatment where the common shock variance is largest, disincentives do occur but not to the degree predicted by our theory and simulations. In addition, in the same treatment, no disincentive effect was observed under the absolute performance contract so there appears to be some support for our theory, at least in a qualitative sense. One possible explanation for why disincentive effects are not as strong as predicted is that some people may care about outperforming others, as suggested by van Dijk, Sonnemans, and van Winden, which can moderate the disincentive effect. We also find that, under tournaments, earnings variability experienced by our subjects is decreasing in the variance of the common shock. This is consistent with the prediction that tournaments are effective at reducing earnings risk in the presence of important common shocks. Finally, a byproduct of our analysis is that we can make inferences about earnings inequality under tournaments and absolute performance contracts. Our experimental results showed that the disparity in earnings between low ability and high ability subjects is greater under tournaments than under absolute performance contracts. This might be a concern to contract administrators who are worried about equity, morale issues, or social preferences (Kennedy 1995; Charness and Kuhn 2004; Bandiera, Barankay, Rasul 2005). In addition, Lazear (1989) suggests that earnings equality can possibly suppress uncooperative behavior when workers can affect each others output.

I. Overview of the Experiments

Prior to describing our theory and providing the details of our experiments, we will provide an overview of our experiments. We conducted three experiments using undergraduate students at a major university in the Midwest. For each experiment, we recruited twelve subjects via posters and email lists across a number of departments on campus. These subjects arrived in a room and were randomly assigned to twelve chairs. For tournament sessions, subjects were randomly matched together to form pairs, but subjects were not informed of the personal identity of the other pair member. Each subject was informed about whether her effort cost function is "high" or "low" and she knew that if she was a high cost subject (low ability subject), her pair member was a low cost subject (high ability subject) and vice versa. Every subject was informed about her cost function as well as her pair member's cost function so that all cost functions were common knowledge.

Every subject was told that she can earn money by making good decisions during the experiments. Each experiment contained four sessions of ten rounds each, where the first session of the night was a tournament session, followed by an absolute performance session. After the first two sessions were completed, we conducted another tournament and another absolute performance session. However, subjects did not gain automatic entry into the second half sessions; instead, they had to bid their way into these sessions through an auction using their experimental earnings from the first two sessions. The sum of a subject's bids could not exceed that subject's accumulated earnings from the initial sessions. The ten highest bidders for the tournament session got to participate in the post-auction tournament. Similarly, the ten highest bidders for the absolute performance session got to participate in the additional absolute performance session. To maximize subject take-home pay, we required that the ten highest bidders pay only the amount offered by the tenth place bidder. Our auction design should, in principle, induce subjects to shade their bids below their true valuation so that subjects would not give up too much of their earnings. Underbidding is not a problem for us because we are not interested in inducing truthful revelation; instead, we are interesting in inducing self sorting of subjects into the second half experiments, as we will discuss in the next paragraph. This sorting requires only an ordinal ranking of bids, and since all subjects followed the same auction rules at the same time and location, all subjects should underbid by the same amount thereby preserving ordinal rankings.

The auction helped eliminate unmotivated subjects and also induced some sorting of subjects into the second half sessions based on motivation, preferences, beliefs and skills. This type of sorting is a feature of competitive real world markets where certain types of people self select themselves into certain occupations and only remain in these occupations if they are sufficiently competent. Indeed, the recent experimental literature has begun to tackle problems associated with sorting or self selection (e.g. Lazear, Malmendier, and Weber 2006). Moreover, the repeated sessions helped to moderate learning effects. By the time subjects reached the second half sessions, they had become "experienced" subjects, i.e., learning effects within these second half sessions should be minimal. Thus, we expect the post-auction results to be relatively free of learning effects and contamination caused by non-self selection.

Once all four sessions were completed, subjects filled out an exit questionnaire and were paid in cash for their performance for the evening. In the next section, we discuss the underlying theory motivating our experiments and the experimental parameters employed.

II. Theory, Experimental Parameters, and Predictions

A. Theory

The underlying theory motivating our experiments has been discussed in other papers on experimental tournaments (e.g. Bull, Schotter and Weigelt 1987 (BSW); and Schotter and Weigelt 1992 (SW)), and is motivated by the work of Lazear and Rosen (1981) and Green and Stokey (1984). A key difference between our experiments and experiments from previous studies is that we include a common shock. We provide a brief discussion of the theory in this section.

Consider a two-player tournament where each subject chooses a costly, non-contractible effort denoted by, e_i, for i =1, 2. Performance for subject i is stochastically related to effort via the production function:

(1) $y_i = e_i + u_C + u_i$ $\quad\quad i$ =1, 2

where y_i is performance, u_C is a common shock, and u_i is an idiosyncratic shock that is independently and identically distributed across agents. Random variables are distributed as follows: $u_C \sim N(0,\sigma_C^2)$, $u_i \sim N(0,\sigma^2)$, $Cov(u_C,u_i)=0$, and $Cov(u_i,u_j)=0$, $\forall\ i \neq j$.

The tournament compensation rule is simple: if $y_i > y_j$, then player i receives a high payment denoted by R and player j receives a low payment, r, where $R > r$, and vice versa.[4] Moreover, both agents face effort-cost functions satisfying: $c(0)=0$, $c'(e_i)>0$ and $c''(e_i)>0$. We also adopt the cost structure used by SW in their experiments, which is of the form,

(2) $c_i(e_i) = \dfrac{\alpha_i e_i^2}{k}$

where $k > 0$. SW allowed for cost (ability) heterogeneity by letting α_i vary across agents. We let $\alpha_i = 1.5$ for half the agents and $\alpha_j = 1$ for the other half - a high-cost agent is always matched with a low-cost agent.

Under tournaments, the probability of agent i receiving the high payment is Prob($u_i - u_j > e_j - e_i$) where $u_i - u_j \sim N(0, 2\sigma^2)$. When common shock variance is the majority of total variance, then $\sigma_C^2 + \sigma^2 > 2\sigma^2$ and the *total* variance is lower under tournaments than under absolute performance contracts because σ_C^2 is eliminated under tournaments in exchange

[4] In the case of a tie, the high cost agent was declared the winner.

for the smaller σ^2 under absolute performance contracts. If we let Prob($u_i - u_j > e_j - e_i$) = 1 - $F(e_j - e_i)$ where $F(\bullet)$ is the CDF of $u_i - u_j$, agent *i's* objective function becomes,

(3) $$E(\pi_i^T) = \left[1 - F(e_j - e_i)\right]R + F(e_j - e_i)r - \frac{\alpha_i e_i^2}{k}$$

which can be written as:

(4) $$E(\pi_i^T) = r + \left[1 - F(e_j - e_i)\right]\left[R - r\right] - \frac{\alpha_i e_i^2}{k}$$

Similarly, agent *j's* objective function is:

(5) $$E(\pi_j^T) = r + F(e_j - e_i)\left[R - r\right] - \frac{\alpha_j e_j^2}{k}$$

The two agents essentially play a game where effort choices are strategies and earnings are given by (4) and (5). First-order conditions are,

(6) $$\frac{\partial E(\pi_i^T)}{\partial e_i} = f(e_j - e_i)\left[R - r\right] - \frac{2\alpha_i e_i}{k} = 0$$

(7) $$\frac{\partial E(\pi_j^T)}{\partial e_j} = f(e_j - e_i)\left[R - r\right] - \frac{2\alpha_j e_j}{k} = 0$$

where *f*($\bullet$) is the density function. If cost functions were identical ($\alpha_i = \alpha_j = 1$), the solution is straightforward. Conditions (6) and (7) suggest that,

(8) $$\frac{2e_i}{k} = f(e_j - e_i)\left[R - r\right] = \frac{2e_j}{k}$$

so that $e_i = e_j = e^*$ which is a symmetric Nash equilibrium. It is clear that the density function *f*($\bullet$) will be evaluated at zero so that $f(0) = \frac{1}{\sqrt{2\pi(2\sigma^2)}}$ giving us,

(9) $$e_i = e_j = e^* = \frac{k\left[R - r\right]}{2\sqrt{4\pi\sigma^2}}$$

However, our study features equilibrium effort levels that differ between high and low cost agents; hence, the normal distribution is not evaluated at $f(0) = \frac{1}{\sqrt{2\pi(2\sigma^2)}}$ making the first order conditions much more difficult to solve. We choose specific parameter values for our model (discussed in sub-section B) and use numerical methods to solve the first-order conditions. Like previous studies, we restrict subjects' effort choices to be an integer from zero to 100; i.e. $e_i \in \{0,1,...,100\}$

Before we discuss parameters and equilibrium predictions, we briefly discuss the model used to motivate the absolute performance contract experiments. Our absolute performance contract is similar to the tournament except agent *i* receives the high payment *R* if output exceeds a fixed standard *y** and *r* otherwise. The probability that agent *i* receives the high payment is Prob($y_i > y^*$) = Prob($u_C + u_i > y^* - e_i$), where $u_C + u_i \sim N(0, \sigma_C^2 + \sigma^2)$. Letting $G(\bullet)$ be the CDF of $u_C + u_i$, we have Prob($u_C + u_i > y^* - e_i$) = 1 - $G(y^* - e_i)$. Agent *i's* objective function is,

(10) $$E(\pi_i^F) = r + [1 - G(y^* - e_i)][R - r] - \frac{\alpha_i e_i^2}{k}$$

with a first-order condition of

(11) $$\frac{\partial E(\pi_i^F)}{\partial e_i} = g(y^* - e_i)[R - r] - \frac{2\alpha_i e_i}{k} = 0$$

Numerical methods are also used to solve (11). In the next section, we discuss our choice of parameters to generate our solutions and predictions.

B. Predictions

So far, we have discussed our model and theory under the assumption that agents are risk-neutral. While Green and Stokey suggest that risk aversion is important for elevating the importance of tournaments in the presence of common shocks, it would be extremely difficult to identify experimental subjects' risk preferences. Moreover, the type (e.g. CARA, CRRA, prospect theoretic preferences, etc.) and the level of risk aversion are likely to be heterogeneous across subjects so that there may be a plethora of possible equilibrium effort levels. Thus, any risk preference structure we impose on subjects is likely to be no less arbitrary than assuming risk neutrality. While we recognize the drawbacks from assuming risk neutrality, we maintain this assumption because it keeps the analysis manageable. At the same time, risk neutral predictions are useful for heuristic purposes and deliver a useful starting point for generating qualitative hypotheses.

We now discuss our process of choosing experimental parameters, which will determine our equilibrium predictions. Like BSW, we choose our cost function parameter, k to be 10,000. We also assume that total variance, $\sigma_C^2 + \sigma^2 = 500$, but we allow the ratio $\frac{\sigma_C^2}{\sigma^2}$ to vary across experiments, which allows us to explore the effectiveness of tournaments at reducing payment variation in different common shock environments. The columns in table 1 represent the different relative magnitudes of σ^2 and σ_C^2 in our three experiments.

In choosing the payments R and r for the tournament, note that firms are typically interesting in inducing a specific level of effort from their workers, at least in the short run. Thus, they typically design an incentive scheme to implement an effort profile. Implementing a particular effort profile involves choosing the *spread* between R and r to ensure incentive compatibility as dictated by equations (6), (7) (for tournaments) and (11) (for absolute performance contracts). The choice of the *level* of r is used to determine expected earnings for the agents. The average effort level we try to implement is $e^* = 37$ as this did not appear to be an obvious number that subjects spuriously might choose, thereby biasing the results. We therefore had to choose the spread $R - r$ to implement an effort level of 37, on average. Because the effort cost functions differed across agents within a pair, high and low cost agents had different Nash equilibrium effort levels, but the average of the two was 37. The low payment r was used to determine the ex ante earnings for agents. For each experiment, we chose r to ensure that high cost agents receive expected earnings of \$14.20 and low cost agents receive expected earnings of \$23.50 for an average of \$18.90. \$19 is about the going rate for a two-hour student experiment on the host campus. Each experiment consisted of four ten-round sessions for a total of forty rounds of play. Thus, the per-round expected earnings for the high cost and low costs agents were \$0.36 and \$0.59, respectively, for an average of approximately \$0.47.[5]

Once the targeted effort levels and expected earnings were identified, we pin down optimal values of R and r (see row 1, table 1). Using experiment 2 as an example, with a common shock variance of $\sigma_C^2 = 250$, the optimal payment spread is $R - r = 0.62$. In principle, this spread should induce high cost subjects to choose e = 30 and low cost subjects to choose e = 44 for an average of e = 37 (see row 2, table 1). To pin down r, numeric simulations showed that r = 0.33 would provide high and low cost agents with per round expected earnings of approximately \$0.36 and \$0.59, respectively (see row 4, table 1).[6] Parameters for the other experiments, where we varied the relative sizes of the common and idiosyncratic shocks, were generated in a similar fashion.

In calibrating parameters for the absolute performance contract, we wanted the average effort level to be 37 as under the tournaments. This allowed us to study how absolute

[5] Although there are some minor rounding errors.

[6] We say "approximately" because our numerical calculations had minor rounding errors. For example, effort for a high cost agent was actually 29.98 for an idiosyncratic variance of 250 and a pay spread of 0.62. The expected earnings were also slightly different from \$0.36 and \$0.59 due to minor approximation errors.

performance contracts were relative to tournaments in achieving the same performance objectives. Second, we wanted to maintain the same expected earnings for the agents, which is to say that the firm would have to satisfy the same participation contracts under both contracts.

Note that the incentive compatibility constraint for the agent is given by (11) so that if we wanted an average effort level of e = 37, we can solve for the payment spread that would induce this average effort level.[7] However, we had to first choose a fixed standard y^*, which output must exceed in order for the agent to receive the high payment R. An obvious choice was y^* = 37, but we avoided this choice because we did not want to provide our subjects with a focal point so that they might naturally gravitate toward a solution of 37. Instead, we chose $y^* = 41$ and then adjusted our payment spread to ensure that 37 was the optimal choice. Our numeric simulations yielded an optimal wage spread of $R - r = 0.55$ (row 1, table 1). While this spread produced an *average* of e = 37, we note that the high cost agents optimal effort is e = 27 while the optimal low cost effort is e = 47 (row 2). Additionally, the value of r that would result in average expected earnings of \$0.47 per round to satisfy the "participation constraint" was r = 0.40 (row 1). However, while this value of r is consistent with an expected per round earnings approximately \$0.47 when we pool all the agents, the expected per round earnings for the high and low cost agents are \$0.44 and \$0.51, respectively. Hence, the gap in expected earnings between high and low cost agents is lower under the absolute performance contract so that we might expect greater inequality under tournaments.

C. Additional Details of the Experiments

We now provide additional details concerning the implementation of our experiments using experiment 2 as an example. For this experiment, each agent's output is the sum of an effort

[7] We also evaluated the second order conditions to ensure that we are at a maximum.

integer from 0 to 100, an idiosyncratic shock, u_i , distributed $u_i \sim N(0, 250)$, and an aggregate shock, u_C, distributed $u_C \sim N(0, 250)$, to get $y_i = e_i + u_C + u_i$. The output for agent j is similarly defined. We approximated a normal distribution with mean 0 and variance 250 using 300 pennies in a bucket where each penny was marked with an outcome for the random shocks. The outcomes were represented by integers and the frequency for each outcome was determined by approximating the number of outcomes out of 300 that might occur under a normal distribution.[8] Distributions for other values of σ_C^2 and σ^2 (i.e. for experiments 1 and 3) were approximated using the same method. For tournament sessions, if $y_i > y_j$, then agent i gets $R = 0.95$ and agent j gets $r = 0.33$, and if $y_i < y_j$, then agent i gets $r = 0.33$ and agent j gets 0.95. This rule applied for experiments 1 and 3 as well, except that the numeric values of R and r were different. Each low cost subject was matched with a high cost subject in a tournament round and both players were informed about both her cost function and her opponent's cost function. The absolute performance sessions were similar to the tournament sessions, except that each subject played against a fixed standard of $y^* = 41$, rather than against a pair member.

In each session, subjects play ten identical rounds of the game. In each round subjects choose "decision numbers" (effort) from 0 to 100 and enter the decision numbers into individually maintained worksheets as an experimenter records the decisions in a computer. Then one subject draws a "common shock" number from a bucket with frequencies that approximated a normal distribution and all subjects add this number to their decision numbers. Next, each subject draws a number from another bucket with frequencies approximating another normal distribution, and then this individual number was added to the decision number and the

[8] The exact method that we used was to calculate the probability mass function in Excel for a normal distribution with mean zero, and standard deviation 15.8. We then multiplied the probability for each outcome by 300 and rounded it to the nearest integer. The resulting integer represented the frequency for that particular outcome.

common shock number. Copies of the probability distributions for both the idiosyncratic and common shocks were given to subjects prior to the beginning of the experiment and explained in detail so these distributions were common knowledge. The sum of the decision number, the common shock, and the idiosyncratic shock is "performance" (y_i).

In tournament sessions, the administrator would compare outputs of pair members and inform all subjects of the relevant payment received for the round (R or r). Each subject is told only his/her own payment and not the difference in output.[9] For absolute performance sessions, output is compared to the fixed standard of 41. Each subject records the payment in the worksheet and subtracts the decision cost to get net earnings for that round. This pattern is identically repeated in all 10 rounds. At the end of the tenth round, subjects calculated cumulative earnings for the ten rounds and confirmed these figures with the computer maintained records kept by the administrator.[10]

All subjects received cost sheets that mimicked their cost functions, knew the distribution of the numbers in the buckets, and were informed of all other experimental parameters, including opponents' cost functions. Only the personal identity of the pair members was not common knowledge. A session typically lasted between 20-25 minutes; two non-paying, practice rounds were played before each session to ensure that subjects understood the experiment. Complete instructions for the experiments are in the appendix.

A potential criticism of our experimental design is that the order of our sessions is not counterbalanced; that is, the tournament is always conducted first and could give rise to order

[9] This is consistent with the way many comparative performance contracts work where workers are informed about their rankings but are not provided detailed information about competitors' performance.
[10] While the tournament was repeated over 10 rounds, the theory is based on a static model. Such repetition is common in experimental practice because subjects make complex decisions. Moreover, the only subgame perfect Nash equilibrium to a finitely repeated game involves the choice of the Nash equilibrium decision level to the one-shot game. Thus, predictions concerning equilibrium play were independent of finite repetition (BSW).

effects. This was done to minimize the potential for subjects to use the fixed standard of 41 as a focal point for choosing strategies in a tournament setting. Moreover, this issue is not as problematic as it might be in some experiments because we will primarily focus on the data for both contract types after the auction.

III. Analysis and Results

A. Hypotheses

Prior to analyzing our results, we state the primary hypotheses that we are interested in for this study. First, our simulations predict that the probability of "winning," (i.e. receiving the high payment *R*) for high cost subjects decreases as σ_C^2 increases and the opposite is true for low cost subjects. In fact, one can see from experiment 3, where σ_C^2 is largest, that the probability of the high cost subject winning is only 20% (table 1). Note in particular that, if the high cost subject "quits" and exerts no effort, she still has an 8% (in parentheses) chance of winning, which is only 12% lower than exerting the equilibrium effort of 30 and much less costly. This means that the expected earnings to the high cost agent from exerting no effort is $0.40 (in parentheses) per round whereas it is only $0.33 if equilibrium effort is exerted. If the high cost agent quits, then the optimal response of the low cost agent is to exert an effort level of 24 (in parentheses), which is far below the equilibrium effort of 44. Thus, a disincentive effect is created for both subjects. Experiment 3 provides us with a treatment that allows us to explicitly test for disincentives of mixed tournaments.

Hypothesis 1*A:* When the common shock is important enough (i.e. $\sigma_C^2 = 350$), then the tournament induces a disincentive effect. Specifically, effort for both types of subjects will be lower in experiment 3 ($\sigma_C^2 = 350$) than in experiments 1 and 2 ($\sigma_C^2 = 150/250$).

Hypothesis 1B: Since the absolute performance contract does not eliminate the common shock, effort should remain invariant to increases in the relative size of σ_C^2; i.e. no disincentives effects will occur.

Another prediction from our simulations is that the variance in earnings under tournaments decrease as σ_C^2 increases (see table 1), which is consistent with the well known result that tournaments can reduce earnings risk by filtering out the common shock. Predicted earnings variance under tournaments is larger than earnings variance under absolute performance contracts when σ_C^2 is small to medium (experiments 1 and 2), but the reverse is true when σ_C^2 is large. Note that this prediction is not affected by potential disincentive effects in experiment 3 as earnings variance is lowest in experiment 3 regardless of whether subjects play the non-disincentive equilibrium or the disincentive equilibrium.

Hypothesis 2A: Under tournaments, the variance of earnings to subjects should be decreasing with σ_C^2.

Hypothesis 2B: Under absolute performance contracts, the variance of earnings to subjects should be invariant to σ_C^2.

If we take earnings variance is an approximation of risk, then clearly the tournament is more effective at reducing earnings risk when common shocks are important. Thus, if subjects are risk averse, the principal should be able to either provide stronger incentives under tournaments (relative to absolute performance contracts) or reduce the wage costs of implementing a specific effort profile.[11] Of course, in principle, risk averse agents may care

[11] Note that our pay spreads, $R - r$ were determined under the assumption of risk neutrality. If agents are risk averse, then the pay spreads reported in table 1 would be too large and should be reduced if the principal is behaving optimally. Nonetheless, the contract that induces lower earnings risk would require a smaller reduction in the spread; i.e. stronger incentives can be maintained under the less risky contract when agents are risk averse.

about the entire distribution of earnings and not just the first two moments. Consequently, using only the variance (or standard deviation) of earnings as a measure of risk is exactly true only under certain assumptions, such as when the utility function is quadratic or if earnings are normally distributed. Nonetheless, in the absence of other information about subjects' risk preferences, we use variance is a manageable approximation of risk.[12]

A byproduct of our experimental design is that we can examine questions concerning pay inequality across tournaments and absolute performance contracts.

Hypothesis 3: The *difference* in earnings between low cost agents and high cost agents is greater under tournaments than under absolute performance contracts.

While earnings inequality is not the main focus of this study, our simulations and experiments do allow us to easily test whether inequality is greater under tournaments or absolute performance contracts. The literature on compensation schemes suggests that earnings equality is desirable to suppress uncooperative behavior when workers can affect each others output (Lazear 1989); to increase morale (Kennedy 1995); or for fairness reasons (Charness and Kuhn 2004). Thus, understanding what type of contract is likely to induce greater earnings inequality, holding productivity constant, might be of interest to scholars studying pay-equity issues. Our simulations suggest that earnings per round for low cost agents would be higher under tournaments rather than under absolute performance contracts (see table 1) so that high ability agents may prefer a competitive scheme. Low cost subjects are expected to earn $0.59 per round (or $0.80 under disincentive effects in experiment 3) under tournaments, but only $0.51 per round under absolute performance contracts. High cost subjects are expected to earn more under absolute performance ($0.44) contracts than under tournaments ($0.36). This

[12] In practice, the standard deviation of earnings is frequently invoked to capture risk in finance. Moreover, a normal distribution may be able to approximate the binomially distributed earnings of our subjects since they play a sequence of independent rounds.

suggests that the expected earnings gap between low and high cost subjects is larger under the tournament scheme.

B. Actual versus Predicted Effort

We begin our analysis by comparing subject effort to predicted levels. Because predicted effort levels are generated under the assumption of risk neutrality, we would not be surprised if actual effort differed from predicted effort. Nonetheless, if subjects choose effort levels that do not deviate on an order of magnitude that is large relative to other experimental studies, it may inspire more confidence in our experimental design.

We begin by examining summary statistics for pooled effort levels in table 2. Note that we partition the data into pre- and post-auction sessions for both types of contracts so that we can see how learning effects and sorting of subjects may have affected the results.

For data combined from all three experiments (row 1, table 2), we find average effort under tournaments is 46.7 in pre-auction sessions and a remarkable 37.5 in post auction sessions, which is not significantly different from predicted effort levels using a Wilcoxon Sign-Rank test. The difference between pre- and post-auction results is also significant at the 5% level, suggesting that learning effects and sorting may matter.

The next three rows of table 2 partition the data across experiments with different size common shock variances, though, as discussed earlier, contract parameters are adjusted to ensure the same optimal effort and earnings across experiments. Thus, we should observe the same average effort across experiments. However, there was noticeable variation in average effort across tournament experiments and all but one of the experiments ($\sigma_C^2 = 150$, post-auction) yielded average effort that was significantly different from predicted levels. Note also that pre- and post-auction effort were significantly different in two of the three experiments.

Pooled average effort under absolute performance contracts was slightly above the simulated optimal effort of 37 for pre-auction (43.2) and post-auction (41.8) data. While average effort did vary across experiments, the variation was not as great in the post-auction data, ranging from a low of 39.8 ($\sigma_C^2 = 350$) to a high of 44 ($\sigma_C^2 = 250$). Moreover, there appears to be less disparity between pre- and post-auction results under absolute performance contracts compared to tournaments, which is not surprising as subjects solve a simpler optimization problem and may learn more quickly.

Table 2 also lists summary statistics for low cost and high cost agents. Several points are worth highlighting. First, when the common shock is the largest ($\sigma_C^2 = 350$), post-auction effort under tournaments is lower than in the other two experiments, which is indicative of some disincentive effects. On the other hand, average effort for both cost types is higher than that predicted at the disincentive equilibrium. High cost subjects should exert zero effort, but the average was 24 (significantly different from zero at the 1% level), while low cost subjects should exert 24, but the average was 34.1 (significantly different from 24 at the 1% level).

Second, in the majority of cases under tournaments, pre- and post-auction effort levels differed significantly at the 5% level for both cost types. Performance under absolute performance contracts appears to be more robust to learning effects and sorting. Given the disparity between pre- and post-auction results for tournaments, whenever pre- and post-auction results conflict, we will give preference to post-auction results.

Third, focusing on the post-auction data, high cost subjects appear to play the predicted tournament equilibrium strategies, on average, with the exception of the experiment where $\sigma_C^2 = 350$. In this experiment, average effort is significantly different from the equilibrium strategy at the 10% level and from the disincentive equilibrium of zero effort at the 1% level.

With regard to absolute performance contracts, in the majority of cases, subjects of both cost types deviate from predicted optimal effort.

How did our experimental results align with simulated outcomes compared to other experimental studies on tournaments? Focusing on the post-auction data, the worst performing outcomes relative to predictions is for low cost subjects in the $\sigma_C^2 = 250$ experiment. Low cost subjects exert mean effort of 57.8 whereas the prediction is 44 for a gap of 13.8. Compared to the BSW experiments, which were similar to ours in design, these deviations from predicted effort were reasonable – some of the BSW effort outcomes, which are measured in the same units, deviated by as much as 22 effort units.

Because subjects played multiple rounds in each session, it would be interesting to examine behavioral patterns across rounds. Figure 1 graphs the evolution of average effort for the $\sigma_C^2 = 150$ and $\sigma_C^2 = 250$ experiments (pooled together) and figure 2 graphs the evolution of effort for the $\sigma_C^2 = 350$ experiment. We constructed a separate graph for the $\sigma_C^2 = 350$ experiment because it has a possible disincentive equilibrium, as well as a regular equilibrium. We graph separate paths for each cost type and for pre- and post-auction experiments. Figure 1 shows that average effort in the post-auction sessions is closer to equilibrium predictions for both cost types. For high cost types in particular, average post-auction effort tends to hover around the equilibrium effort of 30 in most rounds, suggesting that either sorting or learning effects improve subjects' strategies. Low cost subjects typically work harder than predicted but in the post-auction sessions average effort is closer to the equilibrium prediction of 44.

Figure 2 presents a similar graph for the $\sigma_C^2 = 350$ experiment, which is characterized by greater variation across rounds for both subject types. The greater variation may stem from reliance on less data (relative to figure 1) and because there are two possible equilibria in this treatment. One discernable pattern is that average effort tends to be lower for both cost types in the post-auction experiment. In fact, for low cost types, average effort even approaches the disincentive equilibrium of 24 by round 9 of the post-auction session. The same cannot be said for high cost types who, as a group, do not appear to ever approach the disincentive equilibrium of 0 effort.

Figure 3 graphs the evolution of average effort under absolute performance contracts for all experiments. For low cost subjects, the pattern of evolution for pre- and post auction data appears to be similar except for the last round where post-auction averages fall below the predicted level of effort. There appears to be slightly more divergence between pre- and post-auction results for high cost subjects. While subjects collectively exert higher effort than predicted, the post-auction outcomes are closer to the predicted effort level of 27.

Overall, our subjects appear to collectively choose effort levels that do not deviate far from predicted effort levels. This is true for both tournaments and absolute performance sessions, which lends confidence in our experimental design. However, pre-auction results appear to be further from predicted effort levels, especially for tournaments. It's possible that the combination of sorting induced by the auction and the moderation of learning effects helped subjects improve their strategies in the post-auction sessions. As such, most of our analysis from this point forward will focus on post-auction results.

C. Analysis of Disincentive Effects

In this section, we will formally test hypotheses 1A and 1B, which are predictions about the disincentive effects of tournaments. The two hypotheses essentially state that disincentive effects should occur under tournaments but not under absolute performance contracts when the variance of the common shock is large enough.

Figure 4 examines the pattern of average pooled effort (average of high and low cost subjects) across rounds. We pooled the data rather than examine average effort of each cost type separately because disincentive effects are expected to occur for both types. The figure indicates a possibility of disincentive effects as average effort for the $\sigma_C^2 = 350$ experiment is lower than the average effort for the $\sigma_C^2 = 150/250$ across all rounds. A Mann-Whitney test confirms that this difference is statistically significant (p-value = 0.00).

Because subjects played multiple rounds in each session, we had many repeated observations by subject. Hence, across-round observations will not be a part of a random sample, but it does allow us to use panel methods to control for unobservable heterogeneity, such as rate of learning, risk tolerances, among other factors, which may affect subjects' effort choices. We ran a random effects regression (with robust standard errors) where the dependent variable is effort, and the explanatory variables were an experiment 3 dummy, which equals "1" if the observation came from the $\sigma_C^2 = 350$ experiment and "0" otherwise, and a "period" variable to control for time trends. Only the post-auction data was used as we deemed this data to be more reliable.

The results of the regression using the tournaments data are reported in column (1) of table 1. Note that, under hypothesis 1A, the coefficient for the experiment 3 dummy should be negative and significant, and indeed, the estimated coefficient is -12.61 and significantly different from zero at the 10% level of significance. While this suggests that the conditional mean of effort is approximately 12.61 lower under experiment 3 than the other experiments, which is suggestive of a disincentive effect, the evidence does not appear to be overwhelming for several reasons. First, the coefficient is only significantly different from zero at the 10% level. Second, while effort is lower in experiment 3, both types of subjects appear to work harder than predicted under the disincentive effects equilibrium. For example, under the disincentive equilibrium, high cost types should quit and exert zero effort, which should also drag down the average effort of low cost subjects, who are predicted to exert effort of 24. Thus, pooled effort should be 12, which is 25 units lower than the non-disincentive pooled effort of 37. Hence, we would expect the coefficient for the experiment 3 dummy to be -25, which is about twice is large in absolute value as our estimated coefficient of -12.61. Also, according to table 2 (last row), high cost subjects "quit" in only 22% of all trades, and average effort is 24, which is almost as high as what is predicted in the non-disincentive equilibrium. We therefore conclude that while there is some evidence of disincentives from mixed tournaments, one should not discount the statement by van Dijk, Sonnemans, and van Winden that people are motivated by non-monetary factors such as the desire to outperform others.

To test hypothesis 1B, we ran the same regression using only the post-auction absolute performance data – results are reported in column 2 of table 1. Note that the estimated coefficient for experiment 3 is not significantly different from zero (p-value was 0.65) so there is no evidence of a disincentive effect under the absolute performance contract. This result is consistent with hypothesis 1B.

To summarize, we find some evidence of disincentive effects under tournaments but not absolute performance contracts, which is qualitatively consistent with our theory and hypotheses. However, the results should be interpreted with caution as the disincentive effect under tournaments was only significant at the 10% level and the magnitude of the disincentive effect is smaller than predicted.

D. Earnings Risk

Hypotheses 2A and 2B are statements about the variance in earnings under the two contracts. As mentioned earlier, we do not have information about subjects' risk preferences so we use earnings variance as a manageable proxy of risk. It is well known from the tournaments literature that one of the benefits of tournaments is that they eliminate common shocks and therefore can reduce earnings risk exposure to agents. This can benefit the contract administrator because higher powered incentives can be provided to agents when risk is reduced.

Table 4 provides summary statistics for earnings per round along with earnings standard deviations in parentheses using the post-auction data. Focusing on the pooled data in the first four rows, it appears that the standard deviation of earnings decreases under tournaments as σ_C^2 increases. When σ_C^2=150, the standard deviation is \$0.32, which drops to \$0.28 for σ_C^2=250, and finally down to \$0.26 when σ_C^2=350.

To formally test the hypothesis, we ran a random effects regression where the dependent variable is earnings deviation from means, which is defined as,

$$(12) \qquad D = \sqrt{\left(\pi_{ikE} - \bar{\pi}_{kE}\right)^2}$$

where *k* is an index for the contract type (tournament or absolute performance), *E* is an index for the experiment (σ_C^2 = 150, 250, or 350), and *i* is an index for the *i*th observation from experiment *E* under contract *k*. The independent variables are a tournament dummy which equals "1" if the observation is from a tournament session and "0" otherwise; a variable for the standard deviation of the common shock (σ_C) which can take three values, 12.2, 15.8, and 18.7; an interaction term tournament × σ_C; and a period time trend variable. The results of the regression using the post-auction data are reported in column (1) of table 5.

Testing hypotheses 2A and 2B requires an examination of the incremental impact of σ_C on the dependent variable, *D*. Note that the incremental impact of σ_C on *D* under absolute performance contracts is given by the coefficient for σ_C, which, by hypothesis 2B, should not be significantly different from zero. Indeed, the estimated coefficient is -0.003 and it is not significantly different from zero. Thus, we cannot reject hypothesis 2B. Testing hypothesis 2A involves looking at the sum of the coefficients for σ_C and the interaction term tournament × σ_C. The hypothesis predicts that this sum should be negative – that is, deviations should decrease with σ_C under tournaments. Indeed, the sum of the coefficients is negative (-0.01) and significantly different from zero at the 5% level. Thus, our statistical results are qualitatively consistent with hypotheses 2A and 2B, which implies that, as common shocks become more important, tournaments can reduce earnings deviations.

E. Pay Inequality

An important issue that may concern contract administrators is how the distribution of earnings to workers will be affected by the choice of payment scheme. Hypothesis 3 states that tournaments will induce a greater pay gap between low cost and high cost subjects than absolute performance contracts would. Recall that our simulations predict that low cost subjects would earn around $0.59 per round (see table 1) under tournaments contracts and $0.51 per round under absolute performance contracts. At the same time, high cost subjects are predicted to earn $0.36 per round under tournaments and $0.44 under absolute performance contracts. This suggests that the expected pay difference is greater under tournaments.

In table 4, we see that, for the post-auction data across all experiments, low cost subjects did indeed earn more under tournaments than absolute performance contracts as predicted by the simulations. The difference in earnings ($0.54- $0.46 = $0.08) was very close to the difference predicted in the simulations ($0.59-$0.51 =$0.08). With regard to high cost subjects, our simulations predict that they would earn $0.36 per round under tournaments and $0.44 under absolute performance contracts. Scanning table 4, subjects seem to earn less than what was predicted as the overall post-auction earnings were $0.29 and $0.41 under tournaments and absolute performance contracts, respectively. The difference in earnings ($0.41 - $0.29 = $0.12) appears to be slightly larger than the predicted difference ($0.44 - $0.36 = $0.08). Nonetheless, the empirical results are close enough to the predictions of the simulations that we feel that the theory performed adequately.

We now turn to a formal test of hypotheses 3 using regression results. The regression we run has earnings per round as the dependent variable. The independent variables are a tournament dummy, the common shock standard deviation, σ_C, a low cost dummy which takes the value of "1" if the observation is from a low cost subject and "0" otherwise, an interaction term between the tournament dummy and σ_C, an interaction term between the tournament dummy and low cost dummy, and a period time trend variable. The regression results are reported in column (2) of table 5.

Hypothesis 3, which states that the earnings gap between low and high-cost subjects is greater under tournaments than under absolute performance contracts can be tested with the above regression. Because the low cost agents are predicted to earn more under both types of contracts, we expect the low cost dummy coefficient to be positive and significant. Under hypothesis 3, this earnings disparity is expected to increase even more under tournaments so we expect the low cost × tournament interaction coefficient to be positive and significantly different from zero. The results show that this interaction coefficient is positive (0.19) and significantly different from zero at the 1% level of confidence. Hence, our results are consistent with hypothesis 3, which suggests that tournaments implement greater earnings inequality.

With regard to the magnitudes of our estimates, the coefficient for the low cost dummy is 0.07 (significant at the 10% level), which is the difference in pay between high and low cost agents under absolute performance contracts. This equals the predicted gap of 0.07. The estimated gap under tournaments is simply the sum of the low cost coefficient and the lowcost × T interaction coefficient, which is estimated to be 0.26 and significantly different from zero at 1%. This is fairly close to the predicted gap of 0.23. Consequently, our numeric model does a remarkable job of predicting actual earnings outcomes.

IV. Conclusion

In this paper, we discuss the results of an experimental study that investigates two important hypotheses proposed in the literature on tournaments with agents of mixed ability. Specifically, we test for a hypothesized "disincentives effect" which can occur in tournaments with mixed ability agents. We also test the well known hypothesis that, when common shocks are an important source of risk, tournaments can filter out this common shock and reduce earnings risk to workers. This would allow contract administrators to implement stronger incentives and possibly increase productivity. To our knowledge, these hypotheses have not been explicitly tested in the literature.

Understanding whether mixed tournaments trigger disincentives and/or reduce risk is important because there is a potential tradeoff between risk reduction and disincentives; i.e. greater risk reduction may trigger stronger disincentives. To understand this tradeoff, note that disincentives are possible in tournaments if low ability agents "give-up" and high ability agents shirk because they know they will win no matter what. One way to mitigate this disincentive effect is to *increase* the amount of risk faced by workers. Intuitively, if the relationship between effort and performance is noisy, then low ability workers have a better chance of getting "lucky" and winning the tournament which would restore incentives for low ability agents to try hard and for high ability agents to exert effort to reduce the odds of losing due to bad luck. Thus, if mixed tournaments induce significant disincentives, then we may arrive at the counterintuitive conclusion that tournaments are *less* desirable when common shocks are important.

Our primary findings are that disincentive effects arose in our tournament experiments, although these effects are not as strong as we predicted in our theoretical model and subsequent simulations. It appears that many low ability subjects worked hard even when they had clear monetary incentives not to. We also find that tournaments can be very effective at reducing earnings variation when common shocks are important. Taken together, these results suggest that the benefits of risk reduction from eliminating common shocks might outweigh the disincentive effects of reducing risk. Thus, mixed tournaments might still be adequate incentive mechanisms in environments with significant common shocks. One possible direction of future research is to determine why low ability subjects continue to work hard even when they can earn more by quitting. Perhaps behavioral explanations can be used to explain this phenomenon. Another direction for future research would be to develop tournament models that formalize the tradeoff between risk reduction and disincentives to add more clarity to the problem.

A byproduct of our experimental design is that we are able to test whether greater earnings inequality occurs under tournaments relative to absolute performance contracts. We found that there was greater inequality in earnings under tournaments so that a switch to absolute performance contracts may result in a more equal earnings distribution across heterogeneous workers. If equity or morale issues are important in a firm, or if equity can mitigate uncooperative behavior, as suggested by Lazear (1989), then there are additional downsides to using tournaments.

.

References

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2005. Social Preferences and Response to incentives: Evidence from Personnel Data. *Quarterly Journal of Economics* 120, no. 3: 917-962.

Bull, Clive, Andrew Schotter, and Keith Weigelt. 1987. Tournaments and Piece Rates: An Experimental Study. *Journal of Political Economy* 95, no. 1:1-33.

Chan, William. 1996. External Recruitment versus Internal Promotion. *Journal of Labor Economics* 14, no. 4: 555-570.

Charness, Gary, and Peter Kuhn. 2004. Do Co-Workers' Wages Matter? Theory and Evidence on Wage Secrecy, Wage Compression and Effort. IZA Discussion Paper No. 1417, Bonn, Germany, November.

Green, Jerry R., and Nancy L. Stokey. 1983. A Comparison of Tournaments and Contracts. *Journal of Political Economy* 91, no. 3: 349-364.

Holmstrom, Bengt. 1979. Moral Hazard and Observability. *Bell Journal of Economics* 10, no. 1: 74-91.

Hvide, Hans K. 2002. Tournament Rewards and Risk Taking. *Journal of Labor Economics* 20, no. 4: 877-898.

Kennedy, Peter W. 1995. Performance Pay, Productivity and Morale. *Economic Record* 71, no. 214: 240-247.

Knoeber, Charles R., and Walter N. Thurman. 1994. Testing the Theory of Tournaments: An Empirical Analysis of Broiler Production. *Journal of Labor Economics* 12, no. 2: 155-179.

Lazear, Edward P. 1989. Pay Equality and Industrial Politics. *Journal of Political Economy* 97, no. 3: 561-580.

Lazear, Edward P., Ulrike Malmendier, and Roberto A. Weber. 2006. Sorting in Experiments. National Bureau of Economic Research, Working Paper 12041.

Lazear, Edward P., and Sherwin Rosen. 1981.Rank-Order Tournaments as Optimum Labor Contracts. *Journal of Political Economy* 89, no. 5:841-864.

Nalbantian, Haig R., and Andrew Schotter. 1997. Productivity Under Group Incentives: An Experimental Study. *American Economic Review* 87, no. 3: 314-341.

Schotter, Andrew, and Keith. Weigelt. 1992. Asymmetric Tournaments, Equal Opportunity Laws, and Affirmative Action: Some Experimental Results. *Quarterly Journal of Economics* 107, no. 2: 511-539.

Tsoulouhas, Fanis, Charles Knoeber, and Anup Agrawal. Contests to Become CEO: Incentives, Selection and Handicaps. *Economic Theory*, forthcoming

van Dijk, Frans, Joep Sonnemans, and Frans van Winden. 2001. Incentive Systems in a Real Effort Experiment. *European Economic Review* 45: 187-214.

White, Halbert 1980. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica* 48, no. 4: 817-838.

Table 1. Parameters, Simulations and Predictions

	Experiment 1 σ_C^2=150 σ^2=350		*Experiment 2* σ_C^2=250 σ^2=250		*Experiment 3* σ_C^2=350 σ^2=150	
	Tournament	*Absolute Standard*	*Tournament*	*Absolute Standard*	*Tournament*[1]	*Absolute Standard*
Optimal Payments per Round						
R	\$0.99	\$0.95	\$0.95	\$0.95	\$0.90	\$0.95
r	\$0.29	\$0.40	\$0.33	\$0.40	\$0.35	\$0.40
Effort Implemented						
High Cost Agent	30	27	30	27	30 (0)	27
Low Cost Agent	44	47	44	47	44 (24)	47
Average	37	37	37	37	37 (12)	37
Winning Probability						
High Cost Agent	0.29	0.26	0.25	0.26	0.20 (0.08)	0.26
Low Cost Agent	0.71	0.61	0.75	0.61	0.80 (0.92)	0.61
Ex. Earnings per Round						
High Cost Agent	\$0.36	\$0.44	\$0.36	\$0.44	\$0.33 (\$0.40)	\$0.44
Low Cost Agent	\$0.59	\$0.51	\$0.59	\$0.51	\$0.59 (\$0.80)	\$0.51
Average	\$0.47	\$0.47	\$0.47	\$0.47	\$0.47 (\$0.60)	\$0.47
Variance of Earnings						
High Cost Agent	\$0.10	\$0.058	\$0.070	\$0.058	\$0.048(\$0.023)	\$0.058
Low Cost Agent	\$0.10	\$0.072	\$0.070	\$0.072	\$0.048(\$0.023)	\$0.072
Average	\$0.10	\$0.065	\$0.070	\$0.065	\$0.048(\$0.023)	\$0.065
Firm's Expected Wage Cost Per Worker Per Round	\$0.64	\$0.64	\$0.64	\$0.64	\$0.63 (\$0.63)	\$0.64

[1]Numbers in parentheses in this column represents outcomes if the high cost agent chooses the corner solution and exerts zero effort.

Table 2. Summary Statistics for Effort Levels – Means (Standard Deviations)

	Pre-Auction		*Post-Auction*	
	Tournament	Absolute Performance	Tournament	Absolute Performance
Pooled Effort				
All experiments	**46.7*****	43.2***	**37.5**	41.8***
	(21.8)	(17.6)	(21.4)	(16.4)
σ_c=150 exp.	39.9**	39.3**	37.2	41.5***
	(18.4)	(18.2)	(18.3)	(13.8)
σ_c=250 exp.	**56.8*****	**48.6*****	**46.2*****	**44*****
	(16.8)	(18.5)	(20.1)	(18.4)
σ_c=350 exp.	**43.3*****	41.9***	**29.09*****	39.8***
	(25.5)	(14.5)	(21.6)	(16.4)
Low Cost Effort				
All experiments	**54.1*****	51.3***	**45.9****	49.7***
	(19.3)	(13.5)	(19.3)	(12.6)
σ_c=150 exp.	48.3**	50.4**	45.7	50.3***
	(16.3)	(11.3)	(13.2)	(7.1)
σ_c=250 exp.	**63.7*****	58.7***	**57.8*****	56.6***
	(14.1)	(15.0)	(13.0)	(12.2)
σ_c=350 exp.	**50.4*****	44.9*	**34.1***,$^{+++}$**	43.5
	(22.8)	(10.2)	(22.5)	(13.4)
High Cost Effort				
All experiments	**39.3*****	35.2***	**29.1**	32.7***
	(21.7)	(17.5)	(20.1)	(15.5)
σ_c=150 exp.	31.5	28.2	28.7	32.8***
	(16.7)	(17.0)	(18.8)	(13.2)
% of e= 0 outcomes	10%	20%	14%	2%
σ_c=250 exp.	**50*****	**38.5*****	**34.6**	**31.4*****
	(16.6)	(16.1)	(20.8)	(14.7)
% of e= 0 outcomes	0%	6.7%	24%	8%
σ_c=350 exp.	**36.2**	38.8***	**24***, $^{+++}$	34.3**
	(26.2)	(17.4)	(19.5)	(19.1)
% of e= 0 outcomes	23%	8.3%	22%	10%

Note 1: The means were calculated across players and rounds.
Note 2: *,**,*** signifies that effort levels are significantly different from predicted effort levels at the 10% , 5% and 1% levels of significance, respectively, using the Wilcoxon Sign-Rank test.
Note 3: $^{+,++,+++}$ signifies that effort levels are significantly different from the predicted disequilibrium effort levels at the 10%, 5% and 1% levels of significance, respectively, using the Wilcoxon Sign-Rank test.
Note 4: Bold numbers indicate that pre- and post-auction results (for same contract type) are significantly different at at least the 5% level using a Mann-Whitney test.

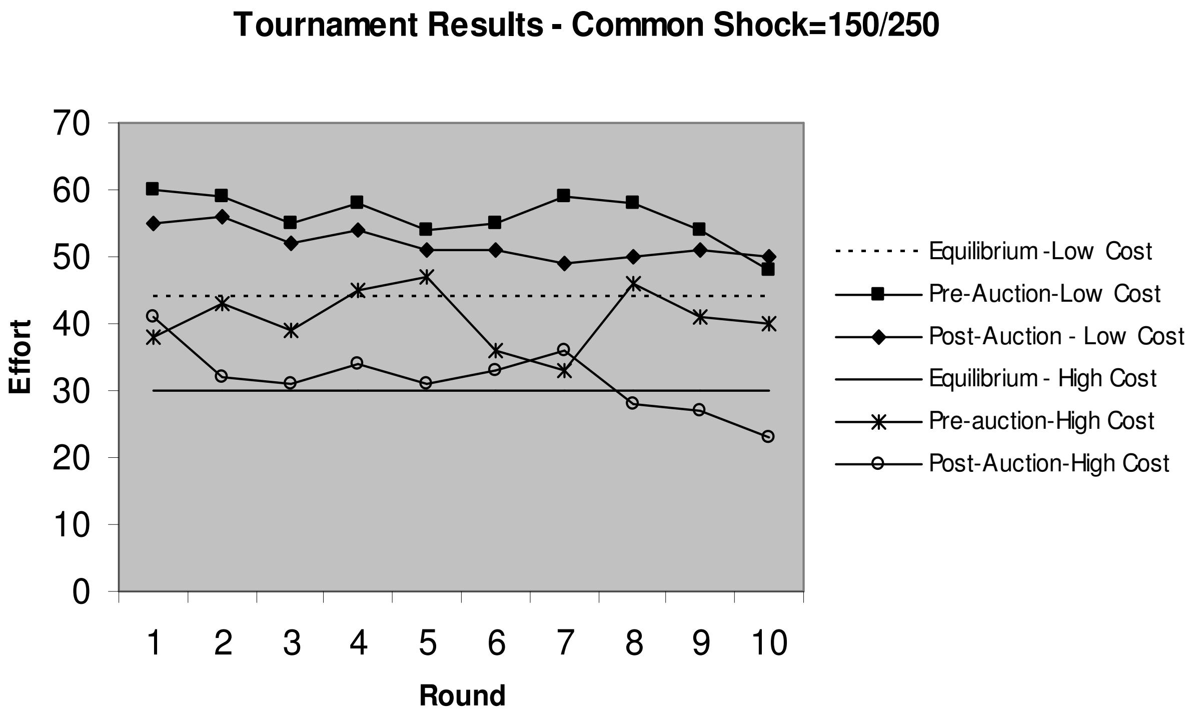

Figure 1. Effort across rounds for the σ_C^2=150 and 250 experiments.

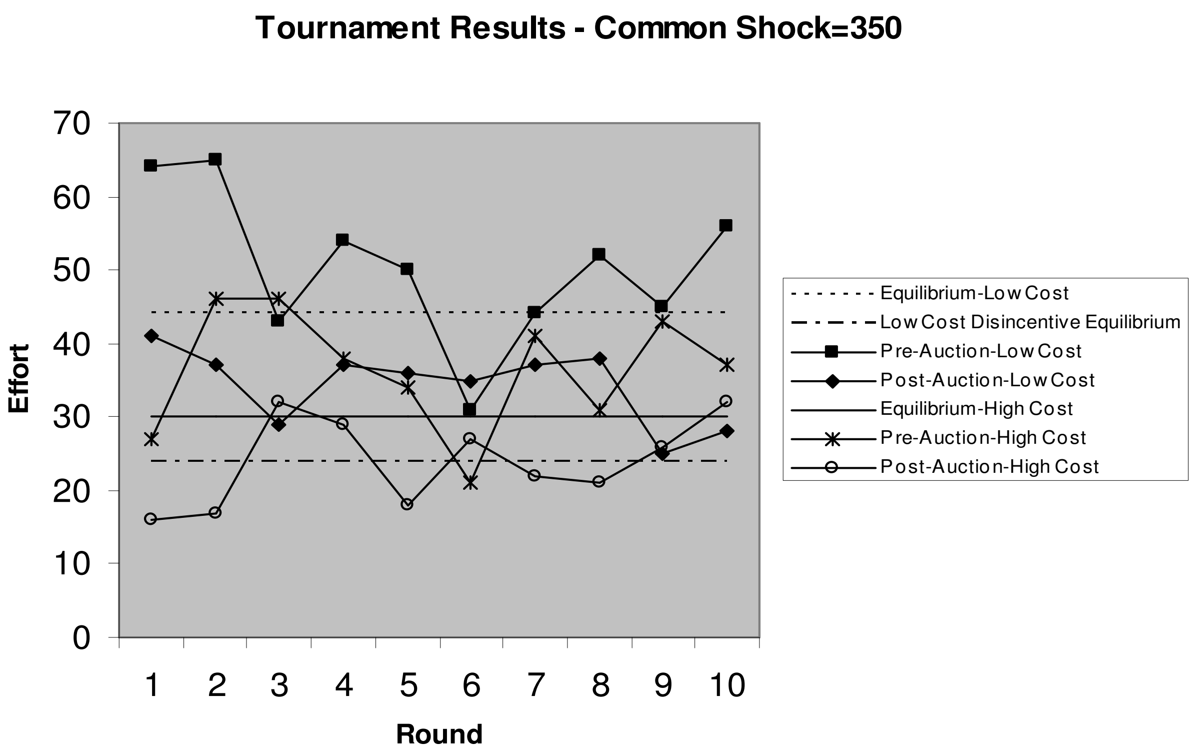

Figure 2. Effort across rounds for the σ_C=350 experiments

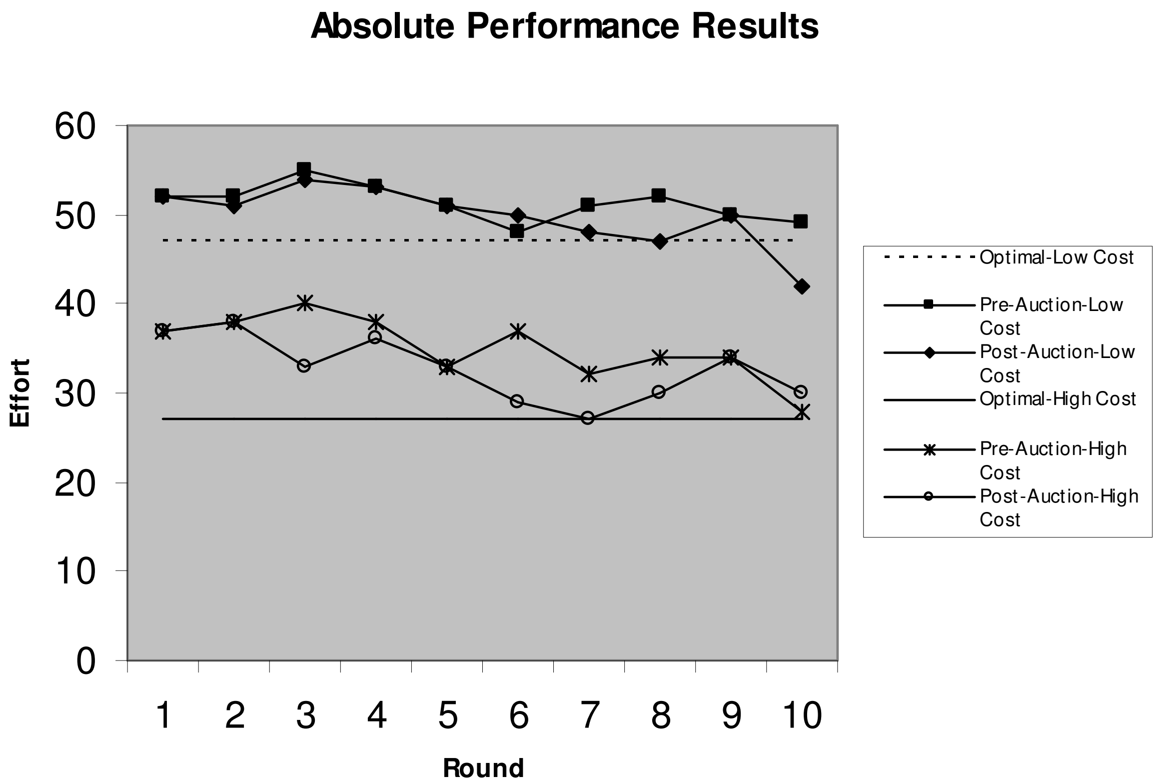

Figure 3. Effort across rounds for all absolute performance experiments.

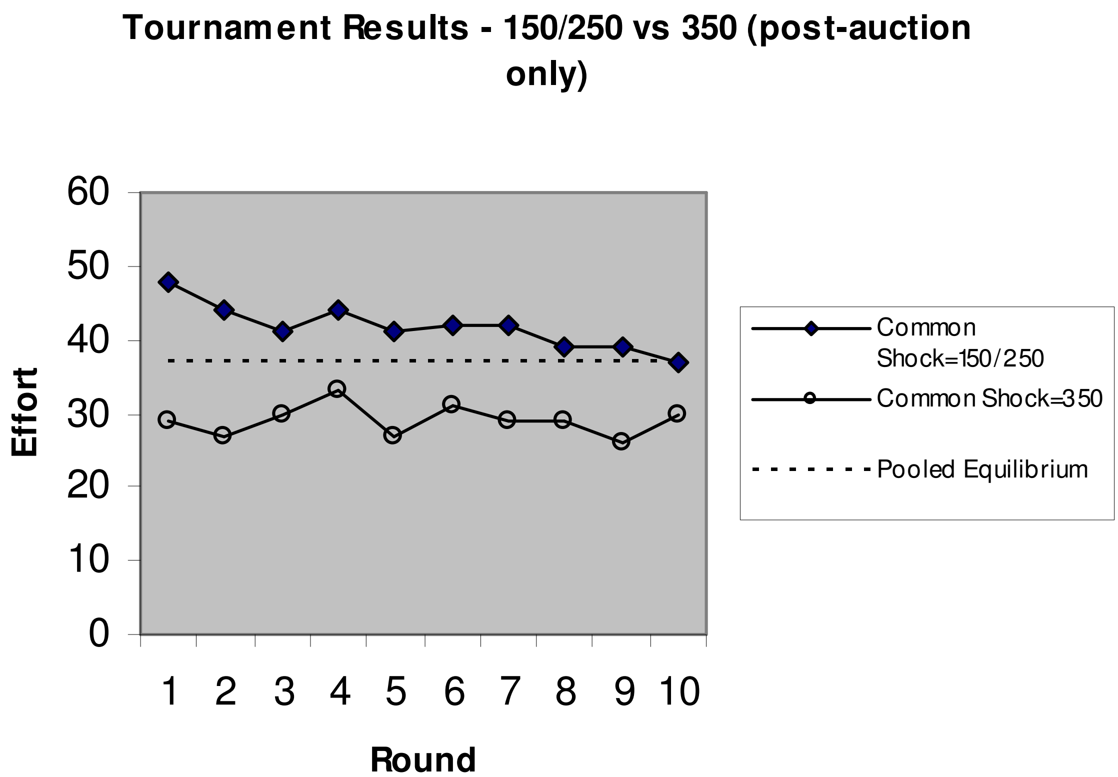

Figure 4. Across round effort comparison between σ_C^2=150/250 and σ_C^2=350 tournament experiments.

Table 3. Effort Regressions (Effort is Dependent Variable)

	Random Effects Estimates	
Data	(1) Tournament Data only	(2) Absolute Performance Data only
Constant	45.25***	47.47***
	(3.79)	(2.74)
Dummy for Experiment 3 σ_c=350	-12.61*	-2.97
	(7.64)	(6.57)
Round trend	-0.65***	-0.86***
	(0.25)	(0.18)
No. of Observations (post auction data only)	300	300

Note 1. *,**,*** signifies that coefficients are significantly different from zero at the 10% , 5% and 1% levels.
Note 2. Robust standard errors are contained in the parentheses below the coefficients and were calculated using the White heteroskedasticity-consistent covariance estimator (White 1980).

Table 4. Summary Statistics for Agents' Per-Round Earnings – Means (Standard Deviations)

	Post-Auction Data	
	Tournament	Absolute Performance
Pooled Earnings		
All experiments	$0.42	$0.44
	($0.29)	($0.25)
σ_c=150 exp.	$0.44	$0.47
	($0.32)	($0.24)
σ_c=250 exp.	$0.34***	$0.49***
	($0.28)	($0.24)
σ_c=350 exp.	$0.47***	$0.35***
	($0.26)	($0.24)
Low Cost Earnings		
All experiments	$0.54***	$0.46***
	($0.25)	($0.24)
σ_c=150 exp.	$0.61**	$0.52**
	($0.28)	($0.24)
σ_c=250 exp.	$0.45	$0.52
	($0.24)	($0.21)
σ_c=350 exp.	$0.57***	$0.37***
	($0.21)	($0.25)
High Cost Earnings		
All experiments	$0.29***	$0.41***
	($0.28)	($0.25)
σ_c=150 exp.	$0.27***	$0.42***
	($0.27)	($0.23)
σ_c=250 exp.	$0.24***	$0.47***
	($0.29)	($0.25)
σ_c=350 exp.	$0.37	$0.31
	($0.26)	($0.23)

Note 1: The means were calculated across players and rounds.
Note 2: *,**,*** signifies that effort levels are significantly different across tournaments and absolute performance contracts at the 10% , 5% and 1% levels of significance, respectively using a Mann-Whitney test.

Table 5. Random Effects Estimates for Earnings (Post-Auction Data)

Variables	**Dependent Variable** (1) Earnings Deviation	(2) Earnings
Constant	0.26***	0.63***
	(0.06)	(0.10)
Tournament dummy (1 if obs. from a tournament session, 0 otherwise)	0.17***	-0.43***
	(0.05)	(0.11)
Common Shock STD (σ_C)	-0.003	-0.02***
	(0.004)	(0.006)
Lowcost dummy (1 if obs. from a low cost subject)	--	0.07*
		(0.035)
Tournament $\times$ σ_C	-0.008**	0.02***
	(0.003)	(0.007)
Tournament $\times$ Lowcost	--	0.19***
		(0.04)
Round Trend	-0.002	0.007**
	(0.002)	(0.0035)
Estimated sum of coefficients for Tournament and Tournament x σ_C	0.16***	--
	(0.05)	
Estimated sum of coefficients for σ_C and Tournament x σ_C	-0.01**	--
	(0.004)	
Estimated sum of the coefficients for Lowcost and Lowcost $\times$ Tournamment	--	0.26***
		(0.04)
Chi-Sq(2) statistic for the joint significance of Tournament and Tournament x σ_C	29.58***	--
Chi-Sq(2) statistic for the joint significance of the σ_C and Tournament$\times\sigma_C$	8.40**	--
Chi-Sq(2) statistic for the joint significance of the Lowcost and Lowcost $\times$ Tournament variables	--	48.17***
No. of Observations	600	600

Note 1. *,**,*** signifies that coefficients are significantly different from zero at the 10% , 5% and 1% levels.
Note 2. Standard errors are contained in the parentheses below the coefficients and were calculated using the White heteroskedasticity-consistent covariance estimator (White 1980)
Note 3. The dependent variables are measured in dollars.
Note 4. Chi-square statistics were reported for the joint tests instead of F-statistics because all that is known about the random-effects estimator is its asymptotic properties. Our regressions were estimated using STATA® which reports Chi-square statistics for random effects regressions.

APPENDIX

Example of Instructions for the Tournament ("Experiment A")

Introduction

This is an experiment about decision making. The instructions are simple, and if you follow them carefully and make good decisions, you could earn a considerable amount of money, which will be paid in cash to you at the end of today's session.

Specific Instructions

As you read these instructions you will be in a room with a number of other subjects. One of these subjects has been chosen to be paired with you by a random drawing of subject numbers conducted before you arrived. This subject will be called your pair member. The identity of your pair member will never be revealed to you and your pair member will never know your identity.

In the experiment you will perform a simple task. In each round of the experimental game you will choose a number between 0 and 100 – this is called your 'Decision Number'. Associated with each Decision Number is decision cost, which is listed in Column B of Table 1. Note that the higher the Decision Number you choose, the higher is the associated decision cost. Also, for each decision number, costs are lower in Table 1A and higher in Table 1B. Whether you have been assigned Table 1A or Table 1B depends on the seat you were assigned. Those that arrived early were randomly assigned to either seat 1, 3, 9, or 2, 4, 10 and will have Table 1A. If you are in seats 5, 7, 11, or 6, 8, 12 you will have Table 1B. If you have Table 1A, then your pair member will have Table 1B and vice versa.

At the beginning of each round of the experimental game you and your pair member will each select a Decision Number separately. Write your number in Column 1 of Sheet 1. Also, record the decision cost associated with your decision number in Column 6 of Sheet 1.

When all subjects have selected their decision numbers, an experimenter will have one subject choose a penny from a bucket with a large number of pennies in it. Each penny in the bucket has a number written on it and the set of all possible numbers range from –35 to + 35. The sheet "Distribution of the Random Number Draw" contains the frequency (number of pennies for each specific number). You will note that more pennies feature numbers closer to zero and the fewer pennies feature numbers close to –35 and +35. In other words, there is a higher probability of drawing numbers closer to zero than numbers far from zero. The penny chosen will be called the 'Group Random Draw Number'. Everyone in the room will enter this number in Column 2.

Then the experimenters will bring buckets around to each of you. You will draw a penny from the bucket and the number on this penny will be called your 'Individual Random Draw Number'. Record your Individual Random Draw Number in column 3 of Sheet 1 and then return the penny to the bucket.

Calculation of Payments

The amount of money you earn in each round will be determined as follows. You will add your Decision Number (column 1) to the Group Random Draw Number (column 2) and to your Individual Random Draw Number (column 3) – write this total in Column 4 of Sheet 1. Your pair member will do the same. The experimenter will also record this information after you receive your Individual Random Draw Number.

Since all subjects have worked in privacy, the experimenter will then compare the totals of you and your pair member. If your total in Column 4 is greater than your pair member, you receive the high payment of $0.95; if your point total is smaller than your pair member, you receive the low payment of $0.33. Whether you receive $0.95 or $0.33 depends only on whether your point total is greater than or less than the point total of your pair member. It does not depend on how much bigger or smaller it is. If there is a tie in total points, the Table 1B pair member gets the high payment.

The experimenter will announce whether you have received a high or low payment. Circle the appropriate payment in Column 5 and subtract the decision cost associated with your decision number, which is in Column 6. Record this difference in Column 7. The amount in Column 7 is your earnings in dollars for the round unless this is a practice round. If this is a paying round, this amount will be added to your running total, which is tabulated in Column 8. Your running total at the end of the 10^{th} paying round is then carried forward to the next experiment.

Before we get started, make sure that you write your chair number on "Sheet 1".

You may also take a minute to look at your pair member's decision cost sheet. Once you have looked at it, please pass it to one of the experimenters and work strictly off of your own decision cost sheet.

Review of Instructions

1. Beginning of Round Announced
2. Choose Decision Number → Record in Column 1
3. Locate associated Decision Cost from Table 1 → Record in Column 6
4. One Subject Draws Group Random Number → Record in Column 2
5. Each subject draw Individual Random Number → Record in Column 3
6. Add Numbers in Columns 1, 2 and 3 → Record in Column 4
7. If your sum is:
   a. Higher than your 'pair member' → Circle $0.95 as your payment
   b. Lower than your 'pair member' → Circle $0.33 as your payment
8. Subtract your 'Decision Cost' from your payment → Record in Column 7
9. If this is a paying round then → Update running total (Col. 8)

Example of Instructions for the Absolute Performance Contract ("Experiment B")

In Experiment A you received the high payment if the sum of your Decision Number, the Group Random Draw Number and your Individual Random Draw Number was greater than your pair member's sum. If your sum was lower than your pair member, you would receive the low payment.

In this Experiment, you will receive a high payment of $0.95 if the sum of your Decision Number, the Group Random Draw Number and your Individual Random Draw Number is greater than or equal to 41. If this sum is less than 41, you will receive a low payment of $0.40. Whether you receive $0.95 or $0.40 as your payment depends only on whether your point total is greater than or equal to 41 – it does not depend on how much bigger or smaller.

All instructions for recording your Decision Number, Decision Cost, Group Random Number, Individual Random Number and payment amount and all instructions for calculating your per round earnings are the same as before.

You will resume tabulating your running total after the one practice round. Please remember to carry forward your net running total from the bottom of Sheet 1 to the top of Column 7 on Sheet 2 so that you can correctly tabulate your running total for this experiment. That is, your running total builds upon your net earnings from the previous experiment and will be carried forward to the next experiment.

Are there any questions?

Review of Instructions

10. Beginning of Round Announced	
11. Choose Decision Number	→ Record in Column 1
12. Locate associated Decision Cost from Table 1	→ Record in Column 6
13. One subject draws Group Random Number	→ Record in Column 2
14. Each subject draw Individual Random Number	→ Record in Column 3
15. Add numbers in Columns 1, 2 and 3	→ Record in Column 4
16. If your sum is:	
a. Greater than or equal to 41	→ Circle $0.95 as your payment
b. Less than 41	→ Circle $0.40 as your payment
17. Subtract your 'Decision Cost' from your payment	→ Record in Column 7
18. If this is a paying round then	→ Update running total (Col. 8)

Table 1A. Decision Numbers and Associated Point Deductions (Seats 1, 2, 3, 4, 9, 10)

Column A: Decision Number	Column B: Decision Cost		Column A: Decision Number	Column B: Decision Cost		Column A: Decision Number	Column B: Decision Cost		Column A: Decision Number	Column B: Decision Cost
0	0.0000		25	0.0625		50	0.2500		75	0.5625
1	0.0001		26	0.0676		51	0.2601		76	0.5776
2	0.0004		27	0.0729		52	0.2704		77	0.5929
3	0.0009		28	0.0784		53	0.2809		78	0.6084
4	0.0016		29	0.0841		54	0.2916		79	0.6241
5	0.0025		30	0.0900		55	0.3025		80	0.6400
6	0.0036		31	0.0961		56	0.3136		81	0.6561
7	0.0049		32	0.1024		57	0.3249		82	0.6724
8	0.0064		33	0.1089		58	0.3364		83	0.6889
9	0.0081		34	0.1156		59	0.3481		84	0.7056
10	0.0100		35	0.1225		60	0.3600		85	0.7225
11	0.0121		36	0.1296		61	0.3721		86	0.7396
12	0.0144		37	0.1369		62	0.3844		87	0.7569
13	0.0169		38	0.1444		63	0.3969		88	0.7744
14	0.0196		39	0.1521		64	0.4096		89	0.7921
15	0.0225		40	0.1600		65	0.4225		90	0.8100
16	0.0256		41	0.1681		66	0.4356		91	0.8281
17	0.0289		42	0.1764		67	0.4489		92	0.8464
18	0.0324		43	0.1849		68	0.4624		93	0.8649
19	0.0361		44	0.1936		69	0.4761		94	0.8836
20	0.0400		45	0.2025		70	0.4900		95	0.9025
21	0.0441		46	0.2116		71	0.5041		96	0.9216
22	0.0484		47	0.2209		72	0.5184		97	0.9409
23	0.0529		48	0.2304		73	0.5329		98	0.9604
24	0.0576		49	0.2401		74	0.5476		99	0.9801
									100	1.0000

Table 1B. Decision Numbers and Associated Point Deductions (5, 6, 7, 8, 11, 12)

Column A: Decision Number	Column B: Decision Cost		Column A: Decision Number	Column B: Decision Cost		Column A: Decision Number	Column B: Decision Cost		Column A: Decision Number	Column B: Decision Cost
0	0.00000		25	0.09375		50	0.37500		75	0.84375
1	0.00015		26	0.10140		51	0.39015		76	0.86640
2	0.00060		27	0.10935		52	0.40560		77	0.88935
3	0.00135		28	0.11760		53	0.42135		78	0.91260
4	0.00240		29	0.12615		54	0.43740		79	0.93615
5	0.00375		30	0.13500		55	0.45375		80	0.96000
6	0.00540		31	0.14415		56	0.47040		81	0.98415
7	0.00735		32	0.15360		57	0.48735		82	1.00860
8	0.00960		33	0.16335		58	0.50460		83	1.03335
9	0.01215		34	0.17340		59	0.52215		84	1.05840
10	0.01500		35	0.18375		60	0.54000		85	1.08375
11	0.01815		36	0.19440		61	0.55815		86	1.10940
12	0.02160		37	0.20535		62	0.57660		87	1.13535
13	0.02535		38	0.21660		63	0.59535		88	1.16160
14	0.02940		39	0.22815		64	0.61440		89	1.18815
15	0.03375		40	0.24000		65	0.63375		90	1.21500
16	0.03840		41	0.25215		66	0.65340		91	1.24215
17	0.04335		42	0.26460		67	0.67335		92	1.26960
18	0.04860		43	0.27735		68	0.69360		93	1.29735
19	0.05415		44	0.29040		69	0.71415		94	1.32540
20	0.06000		45	0.30375		70	0.73500		95	1.35375
21	0.06615		46	0.31740		71	0.75615		96	1.38240
22	0.07260		47	0.33135		72	0.77760		97	1.41135
23	0.07935		48	0.34560		73	0.79935		98	1.44060
24	0.08640		49	0.36015		74	0.82140		99	1.47015
									100	1.50000