TOURNAMENTS, FAIRNESS AND RISK

Revised: August 2005

Abstract

We estimate peoples' willingness to pay to operate under two types of contracts – tournaments and fixed performance standard contracts. Our results are consistent with the notion that subjects have social preferences for fairness and care about risk. That is, when subjects experience greater inequity under tournaments relative to fixed performance contracts, or experience greater revenue risk under tournaments, the gap between willingness to pay for fixed performance and tournament contracts increases, *ceteris paribus.* Our results provide a reason for grower dissatisfaction with tournament compensation schemes independent of possible concerns regarding opportunistic behavior by integrators.

Key words: agricultural contracts, contract regulation, experiments, fairness, social preferences, tournaments.

Relative performance contracts, which reward agents based on performance relative to other agents, can be beneficial to growers in the presence of large common shocks. For example, suppose that bad weather simultaneously reduces the output of all growers in a region. If grower compensation is determined by relative performance rather than absolute performance contracts, then growers are largely insulated from the common shock. In contrast, if the same growers operate under fixed performance standard contracts, then the growers bear the full brunt of the common shock because it becomes more difficult for them to meet the absolute performance standards. Relative performance schemes also have other desirable characteristics which can make them efficiency and welfare enhancing in environments with large common shocks (Hueth and Ligon; Knoeber and Thurman).

Though relative performance contracts are a legitimate incentive and risk management device, some farmers vociferously oppose tournaments, which are a type of relative performance contract (Levy and Vukina; Tsoulouhas and Vukina). Hence, a key puzzle emerges: Why are complaints about tournaments widespread even in industries (e.g. broilers) where common shocks are important?[1] For example, in a web posting, the Rural Advancement Foundation International (RAFI-USA), calls for the banning of "unfair trade practices including tournament pay…"(RAFI-USA). Moreover, The Producer Protection Act of 2000, a model state legislation proposed by several state attorneys general, proposes to ban "unfair practices" such as tournaments.

A plausible explanation for grower opposition to tournaments is that processors, who control inputs and determine compensation, may behave opportunistically (even *illegally*). For example, growers allege that processors manipulate input quality, which can affect performance, or falsify grower rankings to discriminate against certain growers (Hamilton). If these allegations are true, then it is not tournaments per se, but opportunistic behavior that creates problems for growers. In this case, policies that discipline opportunism are preferred to tournament bans, as the latter would only remove an important economic instrument without addressing the core problem.[2] However, if growers dislike tournaments even in the absence of opportunism, then policy makers and processors should know why.

In this article we seek to explain preferences for tournament (T) versus fixed performance standard contracts (F). We use experiments to elicit subjects' *relative* willingness to pay (WTP) for T and F and then examine several factors that explain differences in WTP for F and T, including *profitability*, *riskiness*, and *fairness*. Under standard expected utility theory, profitability and riskiness should be the only factors that influence WTP. However, if subjects have social preferences for fairness, such as inequity aversion (e.g. Fehr and Schmidt), then the degree of inequity induced by contracts might also affect expected utility, which in turn, would influence WTP.

Our empirical results show that differences in WTP for F and T are primarily driven by fairness (inequity) and revenue risk considerations, i.e., subjects were willing to pay more of their own money to participate in contracts that are more fair and induce less revenue variability, *ceteris paribus*. Surprisingly, profitability had little predictive power in explaining WTP differences. Our result that inequity matters provides a possible explanation for why growers often express frustration over tournaments; i.e, growers may perceive tournaments to be less fair and more risky than absolute performance contracts, and that frustration may not be caused exclusively by opportunism.

Our results are broadly consistent with other experimental studies. Economists have shown that, in some settings, subjects do not care solely about monetary payoffs (Camerer and Thaler) and may not operate on the basis of pure self interest (Rabin). Fehr, Gachter and Kirchsteiger, argue that observed contracts are more equitable than those predicted by theory, i.e., principals rarely extract all surplus above an agent's reservation level. When agents get contracts they perceive to be fair, they often reciprocate with effort that exceeds individually optimal levels (Anderhub, Gachter and Konigstein).

The experimental designs of our T and F sessions are identical to the designs of Wu and Roe (2005a), and Wu, et al., which build on the experiments of Bull, Schotter and Weigelt (BSW), and Schotter and Weigelt (SW). During the sessions, subjects make a costly decision that is correlated with performance. However, our experiments include a common shock whereas the BSW and SW experiments do not include this shock. The addition of a common shock is important given their prevalence in agricultural production. In T sessions, compensation depends on performance relative to other subjects. Subjects also participate in F sessions that differ from T sessions only in that payoffs are based on how subjects perform against a fixed performance standard. After these sessions, we held an auction in which subjects can bid into additional sessions of F and T. The bid data contain information about subjects' relative WTP for the contracts.

The Experiments

The purpose of the experiments is to generate bid (WTP) data for T and F contracts using an auction. However, prior to bidding, subjects participate in T and F sessions in order to gain experience operating under the two contracts. Once the subjects gained experience, the auction was held in order to solicit the bids. The highest bidders are allowed to participate in additional T and F sessions in which they can earn additional money.

Over four months, we conducted seven experiments using students at a Midwestern university.[3] Each experiment includes 12 participants recruited via posters or email lists from a variety of departments. At the beginning of an experiment, subjects are randomly assigned to twelve chairs, and then are told they could earn more money beyond their guaranteed $5 show-up fee and that the exact amount depends on decisions made during the experiment. Each subject made approximately $19, on average. Three of the seven experiments involved heterogeneous costs where half the subjects are randomly assigned "high" effort-costs and half the subjects are randomly assigned "low" effort-costs. For the T sessions, a high cost subject is always paired with a low cost subject to form a two-player tournament. For the remaining four experiments, all subjects had symmetric (equal) effort-costs. In the equal cost T sessions, all subjects were randomly paired to form two-player tournaments. No pairing was necessary in the F sessions because subjects only competed against a fixed performance standard and not against another subject.

Each experiment contained four 10-round contracting sessions, where the first session was a T game, followed by an F game. Subjects were then informed that ten of them could participate in another T and another F game where entry to the latter games would be gained via an auction using their earnings from the first two games. Subjects simultaneously submitted sealed bids to enter both additional games. Each subject could bid any dollar amount to enter the second T and F games so long as the sum of the two bids did not exceed total earnings from earlier games. The subjects with the ten highest bids for the second T (F) game gained entry into the additional T (F) game. Participants gaining entry into a second game paid the amount of the tenth place bid.[4] No restrictions were placed on whether a subject could participate in both, neither, or only one of the second games. Finally, subjects completed an exit questionnaire (figure 1) that elicited perceptions about the contracts' relative profitability, fairness, riskiness and fun. Detailed discussions of the experiments, along with experimental instructions, are available in Wu and Roe (2005b).

Predictions

In this section, we will discuss the predictions that we test in the study. The bid data elicited in the auction provides information concerning subjects' relative WTP to operate under two types of contracts. We conduct econometric analysis to determine what factors can explain differences in bids (WTP) for T and F contracts. The factors that we focus on are *earnings* (profitability), *risk*, and *inequity*, and we hope to determine the relative importance of each of these factors in predicting the gap in WTP for T and F contracts. Earnings and risk are important determinants under standard expected utility theory, while inequity can impact utility according to recent theories of inequity aversion.[5]

Standard auction theory predicts that, when subjects value a good more, they will bid more for it. In our experiment, the "goods" are T and F contracts so that subjects should bid more for the contract that yields greater expected utility. We assume that subjects, during the course of playing in the pre-auction contracting sessions, acquire signals (information) that are correlated with the utility they expect to gain from these contracts in the future (the post-auction sessions that they are to bid on). The signals we focus on are earnings (profitability), variance in payoffs (a measure of risk), and degree of inequity (fairness or lack thereof) experienced under the contracts. Maskin and Riley (2000a and 2000b) show that, under weak assumptions, equilibrium bidding is monotonic in signals of the value of a good even in asymmetric auctions with risk-averse bidders.[6] This allows us to characterize the relationship between our factors (signals) and bidding behavior. That is, under expected utility, bids for a contract should be increasing in earnings and decreasing in riskiness (under risk averison). Moreover, when subjects are inequity averse, bids should be decreasing in the amount of inequity experienced. Thus, the primary predictions in this article are that: *subjects should bid more for the contract that induces greater relative earnings, lower relative risk, and lower relative inequity.*

Of secondary interest are predictions about the strategies subjects should play during T and F contracting rounds. If actual strategies chosen by subjects in the contracting sessions are reasonably close to strategies predicted by our economic model and experimental design, then it inspires more confidence in our statistical conclusions about our primary predictions. The main strategy variable in each round of the T and F sessions is "effort." Thus, we discuss predictions about optimal effort that subjects are expected to play in T and F rounds. To describe the strategic role of effort, we will provide an intuitive overview of the experimental designs of the contracting sessions, which are identical to those of Wu and Roe (2005a) and Wu, et al. Readers interested in more formal discussions of the underlying models and/or detailed discussions of the experiments should refer to these articles.

In T sessions, all subjects are randomly paired with another subject to form two-player tournaments. T rules are simple: within each round of a 10 round session, the subject that performs better "wins." That is, if y_i denotes performance for subject *i* and y_j

is performance for subject *j*, then subject *i* wins if $y_i > y_j$ and vice versa.[7] The winner then receives a "high" payment *R* while the loser receives a "low" payment, *r* where *R* > *r*. All subjects were informed of the specific values of *R* and *r* at the beginning of each session. In F sessions each subject is not paired with another subject, but instead competes against some fixed benchmark, y^* (= 41); i.e, each subject *i* receives *R* if $y_i \geq y^*$ and *r* otherwise.

Performance, y_i, is related to effort (the choice variable) via the relationship $y_i = e_i + u_C + u_i$ where e_i is effort, u_C is a random common shock with mean 0, and u_i is a random idiosyncratic shock with mean 0. For the experiment, we restrict e_i to be an integer from 0 to 100. The two random shocks are distributed normally and independently of each other; i.e. $u_C \sim N(0,\sigma_C^2)$, $u_i \sim N(0,\sigma^2)$, $Cov(u_C,u_i)=0$, and $Cov(u_i,u_j)=0$, $\forall\ i \neq j$.[8] Subjects always choose effort prior to the drawing of the random shocks so there is uncertainty about performance at the time effort is chosen. Hence, once effort is chosen, performance variability depends on the variances of u_C and u_i. Thus, increasing effort will increase the probability of good performance but does not guarantee it. However, increasing e_i is costly because we impose an effort-cost function that is increasing and convex in e_i. The specific form of our cost function is given by, $c_i(e_i) = \frac{\alpha_i e_i^2}{10{,}000}$, where $e_i = 0,1,\ldots,100$. The parameter α_i allows us to create cost heterogeneity. In the asymmetric cost experiments, we set $\alpha_i = 1.5$ for high cost subjects, and $\alpha_i = 1$ for low cost subjects. In the symmetric, equal cost experiments, we impose $\alpha_i = 1$ for all subjects. During the experiments, subjects did not have to calculate costs for each effort level because they were supplied with cost tables that mirrored equation (1).[9] Thus, the optimal strategy for a subject is to choose an effort level that will equate marginal expected gains from increasing effort to the marginal cost of increasing effort. This condition is the first order condition for some expected payoff function for each subject, and is analyzed more formally in Wu and Roe (2005a) and Wu, et al.

Optimal effort depends on the experimental parameters that we choose. For instance, the spread between *R* and *r* can influence optimal effort chosen by subjects because it increases the gap in expected payoffs between winning and losing. Therefore, by choosing *R* and *r* appropriately, we can, in principle, induce any equilibrium effort between 0 to 100 that we want so long as we know something about subjects' preferences. Because we do not know subjects' preferences, we begin with the most commonly invoked assumption in the experimental literature (e.g. BSW), that subjects are risk neutral and maximize expected profits.

For each symmetric cost experiment, our goal is to choose *R* and *r* to induce a risk neutral equilibrium effort level of $e_i^* = 37$. In the asymmetric cost experiments, equilibrium effort for low effort-cost subjects will be higher than for high effort-cost subjects. However, we choose *R* and *r* to implement an *average* (average of low and high cost subjects) effort of 37. The number 37 was chosen because it is not an obvious number upon which subjects can anchor. Numerical simulations conducted using Maple software provides specific values of *R* and *r* that induce the desired equilibrium effort.[10]

We ran a total of seven experiments, with three of the experiments involving asymmetric cost subjects where half the subjects were randomly assigned "high" effort-cost tables and the other half "low" effort-cost tables. And then each high cost subject was paired with a low cost subject for the T-sessions. For the F-sessions, no such pairing was necessary because each subject competes only against the fixed threshold y^*(=41). In the four symmetric cost experiments, all subjects were assigned identical effort-cost tables. Columns (5) and (6) of table 1 list the payments R and r generated for the different experiments via numeric computation. These numbers represent the actual payoffs that subjects received for winning and losing under T, or beating or not beating the fixed threshold under F. Experiments 1, 2, 3, and 7 are symmetric cost experiments and the corresponding R and r in columns (5) and (6) should induce subjects to choose e_i = 37 , assuming subjects are risk neutral. Experiments 4, 5, and 6 are asymmetric cost experiments. The associated R and r in columns (5) and (6) should induce high cost subjects to choose $e_i = 27$ while low cost subjects should choose $e_i = 47$ under F, so average effort should be approximately 37. For the T sessions, high cost subjects should choose $e_i = 30$, whereas low cost subjects should choose $e_i = 44$, which also yields an average of 37.[11] These predicted effort levels under risk neutrality are summarized in column 2 of table 2.[12] The payment r can be adjusted up and down to influence the level of expected payoffs to the subjects. Our target is to provide subjects with an expected payoff of $19 per experiment as this was the going rate for a two-hour experiment at the host institution. Each experiment consists of four ten-round sessions (40 total rounds).

In addition to risk neutral predictions, we conducted sensitivity analysis by assuming subjects are inequity averse and/or risk averse. Our goal was to determine the degree to which alternative preference assumptions might affect equilibrium predictions. In order to generate predictions under the assumption that subjects are risk averse, we had to choose a specific preference structure and numeric values for the associated parameters. Results of a study by Holt and Laury guide our preference and parameter choices. The authors conducted a series of experiments in order to estimate risk preferences in experimental settings. The authors show that data patterns can be modeled using a "flexible" expo-power utility function (Saha), and that subjects often display increasing relative, and decreasing absolute risk aversion. Thus, we conducted numeric simulations using an expo-power utility function with all parameters calibrated to Holt and Laury's estimates. In order to generate predictions under the assumption that subjects are inequity averse, we used a model similar to Grund and Sliwka's. We also assumed that subjects are averse to disadvantaged inequity but not advantaged inequity. Disadvantaged inequity is experienced by a subject if her payoffs are less than the payoffs of some relevant comparison group. For example, if subject i receives an amount, x, whereas a competitor receives an amount y where $y > x$, then subject i experiences disadvantaged inequity. Advantaged inequity is defined analogously except a person receives *higher* payoffs than the payoffs of a relevant comparison group. Fehr and Schmidt suggest that disadvantaged inequity may be more important than advantaged inequity, and Charness and Rabin ignore advantaged inequity altogether.

Predictions made under various assumptions about preferences are reported in table 2 in columns (3)-(5). During the simulations, payments R and r are maintained at the levels reported in table 1 so that we can isolate the impact of preferences on optimal effort. Comparing predictions in column 2 to column 3, one can see that, in every case and under both contracts, inequity aversion slightly increases optimal effort relative to the pure risk neutral case. The intuition is that subjects are motivated by both the desire to increase expected payoffs and to avoid inequity. The desire to avoid inequity provides extra motivation to exert effort.

Column (4) reports optimal effort predictions under the assumption that subjects are risk averse but not inequity averse. A direct comparison of the numbers in column (4) to those in column (2) allows us to assess the impact of risk aversion on optimal effort levels.[13] Again, all payments *R* and *r* are held constant at the levels reported in table 1 across the different scenarios. The key points to note are that, (1) symmetric cost optimal effort under both contracts are barely affected by risk aversion; (2) optimal effort for high cost subjects drop by about 5 units under F and between 2 to 4 units under T; and (3) optimal effort for low cost subjects is hardly impacted. We also report the simulations which combine both risk aversion and inequity aversion in column (5) for interested readers. The results are, for the most part, similar to the results in column (3).

Observed Effort versus Predicted Effort

We begin our statistical analysis by determining whether subjects came close to playing predicted effort levels. While this is not the primary focus of this article, it does provide us with a way to check the validity of our experimental design. If we find that subjects play strategies that are close to predicted strategies, or do not deviate on an order of magnitude that is large compared to other experimental studies, then it may inspire more confidence in our statistical conclusions about our primary predictions. Moreover, this exercise also allows us to assess which preference structure made the best predictions.

Average subject effort is reported in column 6 in table 2. We conduct Wilcoxon signed-rank tests to determine whether average effort chosen by subjects differs significantly from predicted effort. Focusing on predictions made under pure risk neutrality (column 2), we see that actual effort is not significantly different from predicted effort in six of ten cases under T. However, predictions made under F did not perform nearly as well – in only one case did actual mean effort not differ significantly from predicted effort. Moreover, introducing risk aversion without inequity aversion (column 4) did not improve predictive power.

The inequity averse models yield predictions that better mirror actual effort. Predicted effort under T was not significantly different from actual effort in four out of the ten cases (column 3). When risk aversion is added (column 5), this improves to five out of ten cases. Turning to F contracts, in the risk neutral-inequity averse case (column 3), five out of the ten predicted effort levels under F were not significantly different from actual effort levels. When risk aversion is added (column 5), this declines to four out of ten cases. Overall, the inequity averse models seem to predict actual effort at almost a 50% success rate, which is quite good compared to many experimental studies involving agents making complex strategic decisions.[14] These results suggest that it may be plausible to model agent behavior using inequity averse preferences.

Analysis of Willingness to Pay

We now turn to the primary goal of this article, which is to econometrically test whether differences in earnings, risk and inequity can explain differences in WTP for the two contracts. We first discuss the data and our econometric specification. Subsequently, we report the econometric results and provide interpretations.

Data and Econometric Specification

Information about WTP is captured by the bid data obtained from the auction described earlier. Bid data were collected for 79 subjects (5 of 84 subjects were repeat participants and excluded). Each subject bid for both contracts, yielding 158 observed bids. Moreover, we obtain measures of earnings, risk, and inequity from the contracting sessions. Earnings are measured as average earnings per-round. Summary statistics for bids and earnings are reported in columns (2)-(4) in table 3. As for our measure of risk, while there is no universal agreement on a best measure of financial risk (Dargahi-Noubary and Smith), we begin by using the most common measure, which is the standard deviation of earnings.

To measure inequity, consider Fehr and Schmidt's assertion that many people care not only about absolute payoffs, but also about payoffs relative to some "reference group." Thus, in order to measure inequity, a reference group must be identified. In a two-player tournament, an agent's competitor serves as a natural reference group. If subject *i* is paired against subject *j* in a two-player tournament, subject *i* would experience disadvantaged inequity if his payoffs are lower than subject *j's*. Thus, we measure disadvantaged inequity for person *i* by defining the variable $DIE_i^T = \max[g_j^T - g_i^T, 0]$ where g_j^T and g_i^T are the total gross payments received by subjects *j* and *i* during the pre-auction session. Advantaged inequity is measured as $AIE_i^T = \max[g_i^T - g_j^T, 0]$. For F, the reference group must be chosen more carefully. The agent might conjecture that, if she was in a T experiment, neither she or a competitor would experience inequity over a ten-round T session if both parties win exactly five times (i.e., total gross payments would be identical). Thus, receiving the high payment five times represents an "equitable" outcome and a reference point. We define $DIE_i^F = \max[\overline{g}^F - g_i^F, 0]$ and $AIE_i^F = \max[g_i^F - \overline{g}^F, 0]$ to proxy inequity under F, where $\overline{g}^F$ denotes the total amount received if the high payment is received in five rounds. Our proxies of inequity are not perfect because they represent inequity in aggregate payments for ten rounds, whereas simulations are based on single round payoffs. Nevertheless, at the time of bidding, subjects had participated in two ten-round sessions so that aggregate results might be superior to a single round measure in capturing the inequity that subjects experienced. Finally, the prediction that greater inequity should occur under T appears to be preserved under our aggregate measures (see columns 5 and 6 in table 3).

To test our main hypotheses, we would like to determine whether differences in WTP for the two contracts are correlated with differences in profits (earnings), risk and inequity induced by the two contracts. To test this, we construct a variable $\Delta bid_i = bid_i(F) - bid_i(T)$, where $bid_i(F)$ is subject *i*'s bid to enter the additional F session and $bid_i(T)$ is subject *i*'s bid to enter the additional T session. If subject *i* has a higher WTP for F relative to T, then she ought to bid more for F and therefore $\Delta bid_i > 0$. On the other hand, if her WTP is higher for T, then she ought to bid more for T in which case $\Delta bid_i < 0$. Then under our primary predictions, if we ran a regression,

$$\Delta bid_i = \beta_0 + \beta_1 \Delta earnings_i + \beta_2 \Delta fairness_i + \beta_3 \Delta risk_i + \varepsilon_i \quad (1)$$

where the regressors are measures of differences in earnings, fairness and risk for the F and T contracts (see discussion below), then we can test our main predictions by examining the sign and significance of the estimates of β_1, β_2, and β_3.

Differences in earnings is defined as $\Delta earnings_i = earnings_i(F) - earnings_i(T)$ where $earnings_i(F)$ is pre-auction earnings under F for subject *i* and $earnings_i(T)$ is the same for T. Because this is a backward-looking measure of earnings for a contract, we also use an alternative measure based on Question 1 of the Exit Questionnaire (figure 1) which is more of a forward looking measure. Answers closer to "5" ("1") suggest that subjects felt that F (T) was more profitable. It could be the case that subjects' *expectations* of the relative profitability of contracts may differ from actual earnings and the responses to this question allow us to check for this. Under either measure, we predict a positive and significant estimate of β_1; that is, holding other variables constant, when subjects experience (or expect) greater earnings under F relative to T, then we should observe higher bids for F relative to T.

For differences in fairness, we use $\Delta DIE_i = DIE_i^T - DIE_i^F$, which represents the difference in disadvantaged inequity experienced by subject *i* under the two contracts in the pre-auction sessions. Positive values imply that subject *i* experienced greater disadvantaged inequity under T. Similarly, we construct $\Delta AIE_i = AIE_i^T - AIE_i^F$ to capture advantaged inequity. We can include both of these variables into regression (1) so that we would have one estimate of β_2 for ΔDIE_i which we will denote as $\hat{\beta}_{2D}$, and another estimate for ΔAIE_i, which we will denote by $\hat{\beta}_{2A}$. Fehr and Schmidt assume that utility loss due to inequity is greater under disadvantaged inequity rather than advantaged inequity. Thus, our prediction is that both $\hat{\beta}_{2D}$ and $\hat{\beta}_{2A}$ will be positive if subjects are inequity averse and that $\hat{\beta}_{2D} > \hat{\beta}_{2A}$. In other words, we predict that when subjects experience greater inequity (whether advantaged or disadvantaged) under T relative to F, this will increase bids for F relative to T. But the impact of disadvantaged inequity will be larger than the impact of advantaged inequity.

For risk differences, we use $\Delta s_i = s_{iF} - s_{iT}$, where s_{iF} and s_{iT} are the standard deviations of per round earnings for subject *i* in the pre-auction sessions. To check for robustness, we also use a measure of risk based on Exit Question 3. Answers closer to "5" ("1") suggest that subjects felt that F (T) was more risky. This measure allows direct exploration of subjects' *perceptions* of the relative riskiness of the two contracts. Our prediction is that the estimates of β_3 will be negative and significantly different from zero. That is, holding other factors constant, when subjects experience (or perceive) F to be riskier than T, we should observe bids for F to *decrease* relative to bids for T. We also use a measure of risk based on the standard deviation of *gross* earnings (revenues) per round, but we will hold off discussion of this until later in the results section.

Because we ran seven experiments, each involving variations in σ_C^2 (common shock variance), *R* (payment for high performance), *r* (payment for low performance), and/or costs, we must control for these factors to test our main hypotheses. Therefore, the simple specification (1) will not allow us to estimate unbiased coefficients. For example, riskiness and inequity would both be affected by experiment-specific factors (levels of *R*, *r*, and σ_C^2, etc) so that if these factors are not held constant, the estimated impact of riskiness and inequity on bids would be biased. In addition, including these additional variables will allow us to improve the efficiency of our estimates. Ideally, we can include a full set of experiment effects (a dummy for each experiment) with the constant omitted to avoid the dummy variable trap.

Because subjects in asymmetric cost experiments are randomly chosen to be either the high or low cost subject in a pairing and each pair of subjects stayed fixed across all rounds of a game, our results could also be biased by the *league composition effect* (Levy and Vukina). The league composition effect predicts that if subjects are heterogeneous and there are fixed leagues (e.g., subjects are paired with the same partner every round), then for long enough sequences of trading periods, F should welfare dominate T, *regardless* of the size of the common shock. The intuition is that when subjects risk being paired with a particularly tough competitor and pairings remain fixed, then subjects may prefer F contracts to avoid "composition risk."[15] Therefore, subjects in heterogeneous cost experiments will have a natural inclination to prefer F. Given that we run experiments both with homogeneous and heterogeneous costs, we can control and test for a league composition effect by including dummy variables for symmetric and asymmetric cost subjects. The intuition is that, all else equal, asymmetric cost subjects should bid more for F than symmetric cost subjects due to the league composition effect.

In order to account for experimental-specific factors and the league composition effect, we could run a regression that includes a full set of seven experiment-specific fixed effects, as these would control for all experiment specific heterogeneity including cost heterogeneity. However, we used the following, equivalent, specification as it allows us to control for the league composition effect *and* to test for it as well. We have:

$$\begin{aligned} \Delta bid_i = {} & \beta_1 \Delta earnings_i + \beta_2 \Delta fairness_i + \beta_3 \Delta risk_i + \beta_4 fun_i + \beta_5 sym_i + \beta_6 asym_i \\ & + \beta_6 effort_i + \theta_2 d_2 + \theta_3 d_3 + \theta_4 d_4 + \theta_6 d_6 + \theta_7 d_7 + \varepsilon_i \end{aligned} \tag{2}$$

where *sym* and *asym* are dummy variables for symmetric and asymmetric cost subjects. The variables d_k are experiment-specific dummies, where k indexes the experiments listed in table 1. We include *sym* and *asym* and only dummies for experiments 2, 3, 4, 6, and 7 (constant is omitted). This regression is identical to one that includes dummies for all seven experiments and omits the *sym* and *asym* dummies. The advantage of our specification is that we can test for the difference in *asym* and *sym* coefficients. A positive and significant difference provides evidence of the league composition effect.

The *fun* variable, represented by the rating of Exit Question 4, is included to control for the fact that some subjects may find one contract to be more entertaining. Failure to include this variable as a control may bias our hypothesis tests. The *effort* variable is measured as the average effort exerted by subject *i* in the pre-auction T and F sessions. This is a potentially important control variable because it captures some unobserved heterogeneity of the subjects. For example, our simulations show that subjects who are inequity averse and/or not very risk averse will tend to exert higher effort under both T and F. Thus, average effort exerted by subjects might contain important information about preferences and other unobserved factors.

Econometric Results

In this subsection, we report our econometric results and discuss whether the results are consistent with our main predictions. That is, we discuss whether differences in bids for F and T can be explained by relative earnings, relative risk and relative inequity. Six variants of (2) are estimated to account for different measures of *earnings* (two measures) and *risk* (three measures) described earlier. Overall, the results (table 4) are relatively consistent across all regressions.

We begin by focusing on whether differences in bids for F and T can be explained by relative earnings. Recall that we have two measures of relative earnings. Our first measure was defined earlier as $\Delta earnings_i$, which is a subject's average earnings under F minus average earnings under T in the pre-auction session. Under our prediction, we expect the regression coefficient corresponding to $\Delta earnings_i$ to be positive because if subjects earn more under F relative to T in the pre-auction session, then we expect subjects to bid more for F relative to T (recall that the dependent variable Δbid_i was defined as a subject's bid for F minus her bid for T). Surprisingly, none of the coefficients corresponding to $\Delta earnings_i$ were significantly different from zero at either the 5% or 10% levels (see regressions 1, 2, and 3), which contradicts our prediction. Our second measure is based on Question 1 of the Exit Questionnaire. Recall that answers closer to "5" ("1") imply that subjects *expect* F(T) to be more profitable. This variable takes values of either 1, 2, 3, 4, or 5, where higher values imply F is more profitable relative to T. Because we predict that when subjects perceive F to be more profitable, they should place higher bids for F (i.e. Δbid_i increases), we expect the corresponding regression coefficient to be positive. However, none of the estimated coefficients were significantly different from zero (see regressions 4, 5, and 6), which again contradicts our prediction. While our findings may seem counter-intuitive, Rabin (pg. 13) states that there is "overwhelming evidence" that people care about their situation relative to some reference point instead of in an absolute sense. If Rabin's assertion is true, then we should not be surprised if absolute earnings do not correlate with WTP, especially after we include our inequity variables as these inequity measures also serve as indicators of relative earnings.

We now discuss whether differences in bids for F and T can be explained by relative riskiness of the contracts. Our first measure of relative risk is Δs_i, which was defined earlier as a subject's standard deviation of earnings under F minus the standard deviation of earnings under T in the pre-auction session. Under our prediction, we expect the regression coefficient corresponding to this measure to be negative because if subjects experience larger standard deviations of earnings under F relative to T (F is more risky), then we expect subjects to bid *less* for F relative to T (recall that the dependent variable, Δbid, is bid for F minus bid for T). The estimated coefficients for Δs_i are -0.49 and -1.01 (see regressions 1 and 4) and neither estimate is significantly different from zero. Thus, these results do not support our prediction. We also used a subjective measure of relative risk, which is based on Exit Question 3 that elicits subjects' *perceptions* of the relative riskiness of the contracts. Recall that answers closer to "5" ("1") imply that subjects perceive F(T) to be more risky. Thus, under our prediction, when responses to this question increase in value (F is perceived riskier), then subjects should place lower bids for F (i.e. the dependent variable Δbid_i decreases). Hence, we expect a negative regression coefficient. The estimated coefficients are indeed negative at -0.34 and -0.39, respectively (see regressions 2 and 5). Moreover, the estimate, -0.39, is significantly different from zero at the 10% level. While the estimate -0.34 was not significantly different from zero at the 10% level, we calculated its *p-value* to be 0.11 so that it was extremely close. Consequently, there is now some support for our hypothesis that subjects will bid more for the contract that induces lower relative risk.

So far, our two measures of relative risk have produced conflicting results. Hence, we explore the risk question further by using an additional measure of relative risk based on *gross earnings* (revenues) per round; i.e. we construct $\Delta g_i = sg_{iF} - sg_{iT}$ where sg_{iF} and sg_{iT} are the standard deviations of *gross* pay for F and T, respectively, in the pre-auction sessions. Like our other measures, we expect the regression coefficient to be negative under our prediction that subjects will bid more for the contract that induces lower relative risk. Notice that the estimated coefficients are -3.49 and -3.38 (see regressions 3 and 6) and both are significantly different from zero at the 10 percent level of significance. These results are fairly consistent with the results obtained from using the subjective measure of risk based on Exit Question 3. It appears that, as subjects experience greater standard deviation of gross pay under F relative to T, they tend to bid less for F relative to T, which is consistent with our prediction. Our results suggest that subjects care about revenue risk but not profit risk, which seems like a behavioral anomaly. However, the tendency for subjects to separate out costs and account only for gross returns is consistent with the notion of *payment decoupling* in the mental accounting literature. Thaler points out that it is possible for "payment decoupling" in the sense that people sometimes overlook costs and focus on benefits (e.g. revenues), especially in cases where costs are paid first and benefits are experienced later. It is possible for "decoupling" to occur in our experiments given the sequence of events in each T and F round. When a round starts, a subject begins by choosing effort, which also determines effort-costs. Then there are several delays as common and idiosyncratic shocks are chosen and encoded. Only then are final outputs calculated and revenues determined. Consequently, there is a time delay between the determination of costs and the determination of revenues, as is the case in many agricultural production settings. This could explain why subjects may have decoupled costs from revenues, and provides an interesting topic for future research.

We now examine our final prediction which is that differences in bids for F and T can be explained by relative inequity induced by the contracts. Recall that our measure of disadvantaged inequity is ΔDIE_i, which was defined earlier as disadvantaged inequity experienced by subject *i* under T minus disadvantaged inequity experienced by the same subject under F in the pre-auction sessions. Under our prediction, we expect the regression coefficient for ΔDIE_i to be positive because if subjects experience greater disadvantaged inequity under T relative to F, then we expect subjects to bid *more* for F relative to T (dependent variable, Δbid, is bid for F minus bid for T). Similarly, we measure advantaged inequity as ΔAIE_i, which was defined as advantaged inequity realized under T minus advantaged inequity realized under F. The coefficient of ΔAIE_i should also be positive but should be smaller in magnitude than the coefficient for ΔDIE_i. This is because Fehr and Schmidt suggest that people care less about inequity when they are ahead (advantaged inequity) than when they are behind (disadvantaged inequity). We find that the coefficients corresponding to disadvantaged inequity (ΔDIE_i) are positive and significant at the 5 percent level across all six regressions, which is consistent with the notion that subjects are inequity averse. In other words, when our subjects experience greater disadvantaged inequity under T relative to F, they appear to bid more for F relative to T, which is consistent with our prediction. However, the coefficients for ΔAIE_i (advantaged inequity) are not significantly different from zero in any regression. Thus, our subjects appear to care greatly about disadvantaged inequity but about advantaged inequity. Overall, our results provide solid evidence that subjects care about disadvantaged inequity and will pay to avoid it. The fact that the coefficients corresponding to disadvantaged inequity were positive and significantly different from zero across all six regressions provides strong evidence that fairness mattered to our subjects in evaluating contracts.

An interesting side result is that the coefficients corresponding to the difference between the *asymmetric* (dummy variable for asymmetric cost subjects) and *symmetric* (dummy variable for symmetric cost subjects) dummies are positive across all six regressions, which is consistent with a league composition effect, but none of these coefficients are significantly different from zero. This suggests that the league composition effect is not strong in our experiments.

Implications and Conclusion

To the best of our knowledge, we provide the first empirical evidence that fairness affects the way people evaluate tournaments. Our measure of fairness is based on the concept of *inequity aversion* and we find that our subjects' willingness to pay to operate under different types of contracts is correlated with the degree of inequity they experience under the contracts. Economists have paid relatively little attention to fairness until recently, but our results suggest that it may be a significant factor in people's choices, and may, in some contexts, be a more significant driver of behavior than absolute monetary incentives.

Our findings may be of interest to contract designers (e.g. processors) because, if fairness enters growers' utility functions, processors may have to pay a "fairness premium" in order to meet growers' reservation utilities. Standard tournament models may not account for the loss in surplus due to inequity aversion. Thus, there may be a tradeoff between the positive (reducing common shocks) and negative (inequity) effects of tournaments, and an optimal contract should carefully balance the two effects.

Our findings can also help policy makers understand why growers may dislike tournaments. Even in the absence of opportunism, people might perceive tournaments to be unfair. This might provide some economic justification for why growers and policy makers often call tournaments an "unfair practice." However, tournaments also have beneficial incentive effects as our results show that inequity averse agents will tend to exert greater effort to avoid disadvantaged inequity. Future research should focus on whether the negative effects of tournaments outweigh the positive effects. Understanding these tradeoffs may provide another perspective on whether proposals to restrict the use of tournaments contracts would be welfare enhancing.

References

Anderhub, V., S. Gachter, and M. Konigstein. 2002. "Efficient Contracting and Fair Play in a Simple Principal-Agent Experiment." *Experimental Economics* 5(1):5-27.

Bull, C., A. Schotter, and K. Weigelt. 1987. "Tournaments and Piece Rates: An Experimental Study." *Journal of Political Economy* 95(1):1-33.

Camerer, C. 2003. *Behavioral Game Theory*. Princeton: Princeton University Press.

Camerer, C., and R. Thaler. 1995. "Ultimatums, Dictators, and Manners." *Journal of Economic Perspectives* 9(2):209-219.

Charness, G.and M. Rabin. 2002. "Understanding Social Preferences with Simple Tests." *Quarterly Journal of Economics* 117: 817-69.

Dargahi-Noubary, G.R., and S. Smith. 2001. "Measuring Investment Risk Based on Tail Thickness." *Review of Quantitative Finance and Accounting* 16(1): 81-94.

Fehr E., S. Gachter, and G. Kirchsteiger. 1997. "Reciprocity as a Contract Enforcement Device: Experimental Evidence." *Econometrica* 65(4):833-60.

Fehr, E., and K.M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *The Quarterly Journal of Economics* 114(3): 817-868.

Grund, C., and D. Sliwka. 2002. "Envy and Compassion in Tournaments." Discussion Paper No. 647, The Institute for the Study of Labor.

Hamilton, N.D. 1995. *A Farmer's Legal Guide to Production Contracts.* Philadelphia, PA: *Farm Journal*.

Holt, C.A., and S.K. Laury. 2003. "Risk Aversion and Incentive Effects." *American Economic Review* 92(5):1644-1655.

Hueth, B., and E. Ligon. 2001. "Agricultural Markets as Relative Performance Evaluation." *American Journal of Agricultural Economics* 83(2):318-328.

Knoeber, C., and W. Thurman. 1989. "A Real Game of Chicken: Contracts, Tournaments, and the Production of Broilers." *J. Law Econ. And Org.* 5(2):271-292.

Leegomonchai, P., and T. Vukina. 2003. "Dynamic Incentives and Agent Discrimination in Broiler Production Tournaments." Working Paper, Dept. of Agricultural and Resource Economics, North Carolina State University.

Levy, A., and T. Vukina. 2004. “The League Composition Effect in Tournaments with Heterogeneous Players: An Empirical Analysis of Broiler Contracts.” *Journal of Labor Economics* 22(2):353-377.

Maskin, E., and J. Riley. 1989. “Optimal Multi-Unit Auctions.” *The Economics of Missing Markets, Information, and Games.* F. Hahn, ed., pp. 312-335. Oxford.:Cambridge University Press.

___. 2000a. “Asymmetric Auctions.” *Review of Economic Studies* 67(3): 413-438.

___. 2000b. “Equilibrium in Sealed High Bid Auctions.” *Review of Economic Studies* 67(3):439-454.

Rabin, M. 1998. “Psychology and Economics.” *Journal of Economic Literature* 36(1):11-46.

Rural Advancement Foundation International (RAFI-USA). “Program on Contract Agricultural Reform.” Available at http://www.rafiusa.org/programs/CONTRACTAG/CCAR.html

Saha, A. 1993. “Expo-Power Utility: A ‘Flexible’ Form for Absolute and Relative Risk Aversion.” *American Journal of Agricultural Economics* 75(3): 905-913.

Schotter, A., and K. Weigelt. 1992. “Asymmetric Tournaments, Equal Opportunity Laws, and Affirmative Action: Some Experimental Results.” *The Quarterly Journal of Economics* 107(2): 511-539.

Thaler, R.H. 1999. “Mental Accounting Matters.” *Journal of Behavioral Decision Making* 12(3):183-206.

Tsoulouhas, T., and T. Vukina. 2001. “Regulating Broiler Contracts: Tournaments versus Fixed Performance Standards.” *American Journal of Agricultural Economics* 83(4):1062-1073.

Wu, S., B. Roe, T. Sporleder, and N. Nazaryan. 2005. “An Experimental Investigation of Alternative Incentive Schemes with Heterogeneous Agents.” Working Paper, Dept. of Ag., Environmental and Development Economics, The Ohio State University.

Wu, S., and B. Roe. 2005a “Behavioral and Welfare Effects of Tournaments and Fixed Performance Contracts: Some Experimental Evidence.” *American Journal of Agricultural Economics* 87(1):130-146.

Wu. S., and B. Roe. 2005b “*AJAE* Appendix: Tournaments, Fairness, and Risk.” Unpublished manuscript.
Available at: http://agecon.lib.umn.edu/

Figure 1. Exit Questionnaire

Chair #: ______________________

• In Game A your payment depended on how your score in points compared to your pair member. (Note: Game A is the tournament)

• In Game B your payment depended on how your score compared to a fixed standard. (Note: Game B is F)

	Game A		About the Same		Game B
1. If each game were played many times under the same conditions as today, which one would earn you more money?	1	2	3	4	5
2. Which game did you think was more fair to all the participants involved?	1	2	3	4	5
3. Which game did you think was more risky?	1	2	3	4	5
4. Which game did you think was more fun to play?	1	2	3	4	5

Table 1. Experimental Parameters for Seven Experiments

Experiment	σ_C^2	σ^2	Effort-Cost	Tournament Payments R r	Fixed Perf. Payments R r
Experiment 1	250	250	Symmetric $\alpha_i = \alpha_j = 1$	\$0.81 \$0.40	\$0.85 \$0.43
Experiment 2	50	450	Symmetric $\alpha_i = \alpha_j = 1$	\$0.88 \$0.33	\$0.85 \$0.43
Experiment 3	350	150	Symmetric $\alpha_i = \alpha_j = 1$	\$0.77 \$0.45	\$0.85 \$0.43
Experiment 4	250	250	Asymmetric $\alpha_i = 1$ $\alpha_j = 1.5$	\$0.95 \$0.33	\$0.95 \$0.40
Experiment 5	350	150	Asymmetric $\alpha_i = 1$ $\alpha_j = 1.5$	\$0.90 \$0.35	\$0.95 \$0.40
Experiment 6	150	350	Asymmetric $\alpha_i = 1$ $\alpha_j = 1.5$	\$0.99 \$0.29	\$0.95 \$0.40
Experiment 7	0	500	Symmetric $\alpha_i = \alpha_j = 1$	\$0.90 \$0.31	\$0.85 \$0.43

Note. The symbol *R* represents the "high payment" a subject receives for winning under T or beating the fixed performance standard under F. The symbol *r* denotes the "low payment" for failing to win under T or beat the fixed performance standard under F.

Table 2. Predicted and Actual Effort Levels – Top Numbers are for Fixed Performance Contracts (F) and Numbers in Parentheses are for Tournaments (T)

Experiment Type	Simulated Effort Levels under Different Preference Assumptions: Risk Neutral	Inequity Averse	Risk Averse	Risk Averse and Inequity Averse	Actual Effort: Mean Effort Observed in Experiments
Symmetric $\sigma_C^2 = 0$	37 (37)	40* (43*)	36 (37)	40* (42*)	36 (40)
Symmetric $\sigma_C^2 = 50$	37 (37*)	40 (42)	36 (36*)	40 (42)	43 (38)
Symmetric $\sigma_C^2 = 250$	37 (37*)	40 (42)	36 (36*)	40 (42)	42 (36)
Symmetric $\sigma_C^2 = 350$	37 (37)	40 (42)	36 (37)	40 (42)	40 (31)
Asymmetric-High Cost $\sigma_C^2 = 150$	27 (30*)	34* (33*)	22 (28*)	33* (30*)	30 (30)
Asymmetric – High Cost $\sigma_C^2 = 250$	27 (30)	34* (33)	22 (27)	33* (29)	35 (43)
Asymmetric – High Cost $\sigma_C^2 = 350$	27 (30*)	34* (32*)	22 (26*)	33 (28*)	37 (31)
Asymmetric-Low Cost $\sigma_C^2 = 150$	47 (44*)	49* (50)	47 (44)	49* (49*)	50 (47)
Asymmetric – Low Cost $\sigma_C^2 = 250$	47 (44)	49 (49)	47 (44)	49 (47)	57 (61)
Asymmetric – Low Cost $\sigma_C^2 = 350$	47* (44*)	49 (48*)	47* (42*)	49 (46*)	44 (43)

Note: An asterisk indicates actual effort level is not significantly different from this simulated effort level at the 5% level of significance using the Wilcoxon signed-rank test. Reported effort levels were rounded to the nearest whole number.

Table 3. Summary Statistics for Bids, Earnings, Standard Deviation of Earnings, and Inequity - Top Numbers are for Fixed Performance Contracts (F) and Numbers in Parentheses are for Tournaments (T)

Subject Types	Mean Bids	Avg. per-round Actual Earnings	Avg. per-round Standard Deviation of Earnings	Avg. per-session Advantaged Inequity	Avg. per-session Disadvantaged Inequity
Symmetric Cost Subjects	$3.71 ($3.48)	$0.48 ($0.44)	$0.19 ($0.24)	$0.54 ($0.74)	$0.25 ($0.68)
High Cost Asymmetric Subjects	$2.75 ($1.93)	$0.40 ($0.26)	$0.24 ($0.29)	$0.37 ($0.26)	$1.10 $(2.15)
Low Cost Asymmetric Subjects	$3.22 ($2.46)	$0.47 ($0.46)	$0.24 ($0.30)	$1.10 ($2.15)	$0.00 ($0.26)
All Asymmetric Cost subjects	$2.98 ($2.20)	$0.44 ($0.36)	$0.25 ($0.31)	$0.73 ($1.20)	$0.55 ($1.20)
All Subjects	$3.38 ($2.90)	$0.46 ($0.41)	$0.22 ($0.28)	$0.63 ($0.95)	$0.39 ($0.92)

Table 4. Regression Estimates – Dependent Variable is Δbid (bid for F minus bid for T in Dollars)

Variables	Regression (1)	Regression (2)	Regression (3)	Regression (4)	Regression (5)	Regression (6)
$\Delta earnings$ (Profit)	-3.50 (2.78)	-3.37 (2.61)	-3.57 (2.74)	--	--	--
Q1 – Expected Profitability	--	--	--	0.002 (0.17)	-0.17 (0.20)	0.007 (0.17)
ΔDIE (Disadvantage Inequity)	0.56** (0.25)	0.53** (0.22)	0.68** (0.27)	0.48** (0.23)	0.45** (0.19)	0.58** (0.25)
ΔAIE (Advantaged Inequity)	-0.23 (0.15)	-0.20 (0.14)	-0.13 (0.12)	-0.11 (0.14)	-0.13 (0.11)	-0.02 (0.13)
Δs_i (Std. Dev. of Profit)	-0.49 (2.15)	--	--	-1.01 (2.36)	--	--
Q3 – Risk Perceptions	--	-0.34 (0.21)	--	--	-0.39* (0.23)	--
Δg_i (St. Dev. Revenue)	--	--	-3.49* (1.95)	--	--	-3.38* (2.03)
Q4 – Fun	0.26* (0.15)	0.18 (0.14)	0.26* (0.15)	0.20 (0.16)	0.17 (0.14)	0.19 (0.16)
Symmetric (Sym Cost Dummy)	-2.27** (0.91)	-1.28 (0.86)	-2.32** (0.91)	-1.76* (0.98)	-0.28 (1.15)	-1.81* (0.98)
Asymmetric (Asym. Cost Dummy)	-1.86* (1.04)	-0.67 (1.01)	-1.95* (1.05)	-1.72 (1.13)	0.01 (1.21)	-1.78 (1.11)
Effort (Avg. Effort)	0.045** (0.02)	0.043** (0.02)	0.04** (0.02)	0.04** (0.02)	0.03** (0.01)	0.04** (0.01)
Dummies for Experiments 2, 3, 4, 6, 7	Controls (not reported)	Controls (not reported)	Controls (not reported)	Controls (not reported)	Controls (not reported)	Controls (not reported)
Asymmetric-- Symmetric	0.41 (0.69)	0.62 (0.65)	0.37 (0.67)	0.05 (0.73)	0.29 (0.64)	0.03 (0.70)
R-squared	0.35	0.40	0.36	0.32	0.38	0.33

Note: One and two asterisks indicate significance at the 10% and 5% levels, respectively. White's robust standard errors are in parentheses.

[1] Using data from the broiler industry, Levy and Vukina determine that the percent of production variance due to common shocks exceeds the variance from all other shocks.

[2] Leegomonchai and Vukina provide a rare empirical investigation of opportunism in agricultural contracting. They find no significant evidence that broiler contractors discriminate across growers of differing ability when allocating heterogeneous inputs.

[3] One might question whether the use of students rather than farmers weakens our results. We regard the use of students as a strength because growers' attitudes toward tournaments may be politicized by recent discussions about the "oppressive" nature of tournaments.

[4] Our auction design can induce subjects to shade bids below true valuation. Nonetheless, underbidding does not affect our results because we are only interested in *relative* WTP for T and F (i.e. subjects bid more for the contract they value more), and not absolute, truthful measures of WTP. Relative WTP should not be affected by our auction rules because the bidding for both F and T took place at the exact same time, at the exact same location, and under the same rules; thus, any valuation bias due to the auction design will affect bids for both contracts in the same way thereby preserving relative WTP. The advantage of our design is that it does not create a situation where subjects have to give away too much of their earlier earnings because they will tend to underbid relative to their valuations.

[5] In Wu and Roe (2005b), we use simple models of inequity aversion by Fehr and Schmidt, and Grund and Sliwka to show how inequity affects incentives and utility of agents. In particular, we show that disadvantaged inequity occurs more frequently under T and thus the expected utility loss is greater under T than F.

[6] Our auction also has the features of a sealed high-bid auction with multiple units, where a seller has more than one unit of a good and each buyer can purchase at most one good. Maskin and Riley (1989) generalized the theory of optimal auctions to the multi-unit case; many insights from the single-unit case apply to the multi-unit case.

[7] Ties are broken by a coin flip.

[8] The way we implement the random shock in the experiments is to approximate each shock using 300 pennies each marked with an outcome. The frequency for each outcome is determined by approximating the number of outcomes out of 300 that occur under a normal distribution. For experiment 1 (see table 1), for example, we calculate the probability function in Excel for a normal distribution with mean zero and standard deviation 15.8 (implies a variance of $\sigma^2 \approx 250$), and multiply the probability for each outcome by 300 and round to the nearest integer. We then place the 300 pennies in a bucket and have subjects draw a penny for each shock. All subjects received copies of the probability distributions for all shocks; these distributions were explained in detail prior to the start of a session.

[9] Moreover, subjects were informed of all experimental parameters including opponents' cost tables. In fact, during asymmetric cost experiments, both the low and high cost tables were projected onto a big screen so that each low cost subject knew that her pair member had the high cost table and vice versa.

[10] Details of all simulations discussed in this section along with MAPLE code are available in Wu and Roe (2005b).

[11] Note that payments necessary to induce equilibrium effort varied across experiments. The reason for this is because we also varied the relative magnitude of the common shock

variance across experiments and had to adjust *R* and *r* accordingly. We varied the common shock variance because this is an important comparative static for assessing tournaments. As was stated in the introduction, a key justification for the use of tournaments is the presence of important common shocks, where "importance" is determined by the size of the common shock variance relative to the idiosyncratic shock variance. The importance of the common shock may affect subjects' attitudes about T and F.

[12] Note that the rows of table 2 are organized differently from table 1 to facilitate comparisons across cost types and common shock sizes.

[13] Note that results would likely differ if we had chosen a different level of risk aversion and/or different risk preferences. However, because our assumptions are derived from Holt and Laury's study, which is a landmark study on risk preferences in experimental settings, we believe that our assumptions are reasonable.

[14] Camerer surveys many experimental studies on game theory and finds that results are typically mixed. In some experiments, standard theory describes behavioral outcomes quite well whereas in other settings it does not. Viewed in this light, our results are not unusual, especially when one considers that our experiments are more complex than some of the games surveyed by Camerer.

[15] We thank an anonymous reviewer for pointing this out to us.