2003 American Agricultural Economics Association Annual Meetings
July 27 – 30, 2003
Montreal, Quebec

PERFORMANCE AND WELFARE EFFECTS OF TOURNAMENT CONTRACTS: SOME EXPERIMENTAL EVIDENCE[1]

Steven Wu and Brian Roe[2]
May, 2003

The Ohio State University
AED Economics
Agricultural Administration Building
2120 Fyffe Road
Columbus, OH 43210

Contact:
wu.412@osu.edu

(614)247-7494

Using experimental economics, we compare the efficiency and welfare effects of tournament and fixed standards contracts. Our findings suggest that economic agents are generally better off under fixed standard contracts unless they face substantial common shocks. Administrators of contracts (principals) also tend to be better off under fixed standard contracts for moderate to small common shocks. Efficiency wise, agents tend to exert higher effort under fixed standard contracts. Moreover, effort under tournaments appears to be declining in the variance of the common shock. Our results suggest that a ban on tournament contracts may generally be better off for both growers and processors except in cases where common shocks are large.

[1] We 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 Tom Sporleder for helpful comments during the research and 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.

[2] The authors are Assistant Professor and Associate Professor, respectively.

I. Introduction

The regulation of agricultural production contracts has been a topic of debate in recent policy discussions at both the state and federal levels. These debates emerged from widespread grower discontent about the fairness of agricultural contracts and have led to a number of legislative proposals that are designed to regulate agricultural contracts. One proposal that has been on the table is the banning of relative performance contracts, which condition growers' payments on the performance of other growers through rank order tournaments. As Tsoulouhas and Vukina (2001, page 1062) stated, "Growers are opposed to a system that bases their payments on how well or how poorly their neighbors perform." Such comparative performance payment systems make it difficult for growers to predict how much they can expect to earn for any given growing period.

While some law makers and grower groups have pushed for the regulation or banning of tournament contracts, relatively few serious economic studies have been undertaken to assess the potential efficiency and welfare effects of alternative payment regimes. A nice study by Tsoulouhas and Vukina (2001) based on analytical modeling suggested that banning tournament contracts and replacing them with fixed standard contracts can increase grower welfare provided that the magnitude of the piece rate in the fixed standard contract is also regulated. Levy and Vukina develop a related theoretical model and derive conditions under which a ban of tournaments would be welfare reducing. Calibrating key model parameters to data from tournaments in the North Carolina broiler chicken sector, they conclude that banning tournaments would be welfare reducing. Roe and Wu apply results from Levy and Vukina's empirical analysis to calibrate an explicitly dynamic model and also find little support that banning tournaments would increase total surplus but that it might benefit low ability producers.

This paper offers experimental evidence about the behavioral and welfare effects of tournaments and fixed standard payment contracts. We followed an experimental approach that was similar to the approaches of Bull, Schotter, and Weigelt (1987) and Schotter and Weigelt (1992). The key difference in experimental design is that our experimental subjects faced both idiosyncratic and common risk whereas these studies exposed subjects only to idiosyncratic risk. We believe that exposing agents to common risk is fundamental in examining behavior under relative performance contracts since a key justification in the literature for having tournaments is to reduce aggregate risk. Indeed, Holmstrom (1982) suggested that if agents' random outputs are not related, then tournaments will only serve to expose agents to more uncertainty without any gain in information about agents' performance.

Our experimental subjects were primarily university students at The Ohio State University. We recruited students and asked them to perform a simple task, which is analogous to effort, under both a fixed standards and tournament contracts. We subsequently calculated their average effort levels, ex-post payoffs, and modeled ex post payoffs for the principal (e.g. the processor).

A brief summary of our experimental findings are, (1) costly effort exerted by subjects is higher, on average, under fixed performance contracts than under tournaments but this difference was mitigated by large common shocks; (2) agent (grower) welfare is higher under fixed standards contracts so long as common shocks are not too important; and (3) for certain types of revenue functions, principal welfare will be higher either when common shocks are of moderate to low importance and they use fixed standard contracts, or when common shocks are very important and they use tournament contracts. These results suggest that, unless common shocks are very important, bans on tournaments may enhance both ex post grower and processor welfare in the short run.[3]

While our experimental subjects are university students rather than agricultural producers, we believe that our results will still reveal useful insights about human behavior under different incentive contracts. Moreover, our results are less likely to be biased by politics since most students are unfamiliar with the controversies surrounding agricultural contract legislation.

In recent years, experimental economics has become an important tool for assessing alternative regulatory regimes. It has been useful in providing insights into number of issues ranging from affirmative action (Schotter and Weigelt, 1992), to alternative mechanisms for public goods (Ledyard, 1994), to evaluation of tradable permits for emissions (Cason and Plott, 1996), among other studies. Eckel and Lutz (2003) suggest that a key advantage of experimental economics is that it can be very useful in evaluating hypothetical situations (e.g. impact of new regulations) that are difficult to replicate in the field.

II. The Theoretical Framework

A. The Tournament Model

The theory of tournaments (e.g. Lazear and Rosen, 1981) states that an agent's payoff depends on how he performs relative to other players rather than against some fixed

[3] While this result appears to be straightforward, one has to remember that tournaments may not get used unless common shocks are fairly sizable in the first place. Hence, it leaves open the question of whether the government should institute a ban given that economic decision makers will naturally adjust their strategies in response to different magnitudes of common shocks.

standard of performance. For simplicity, we will restrict our attention to two player tournaments so that the "winner" receives a high payment in the amount R and the "loser" receives a payment r where $R > r$.[4]

To model behavior under tournaments, consider a risk neutral principal who contracts with two identical risk neutral agents.[5] Following the literature, each agent's output is related in a linear fashion with effort and is defined by the relationship,

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

where e_i is agent *i's* non-contractible effort, u_C is a random variable that is common to both agents in the tournament, and u_i is an idiosyncratic random variable that is independently and identically distributed (i.i.d) across agents. We assume that the random variables are normally distributed, where $u_C \sim N(0, \sigma_C^2)$, $u_i \sim N(0, \sigma^2)$, and $Cov(u_C, u_i) = 0$ for all i.

Compared to other tournament experiments we have referenced, our research is unique in that we include the common shock in the agents' production functions. The inclusion of a common shock is important both at a theoretical and behavior level. On a theoretical level, Holmstrom's (1979) *sufficient statistics theorem* suggests that any information that can improve the principal's ability to infer agents' effort levels ought to be included in the contract; indeed, the sufficient statistic theorem has been frequently used as a justification for the use of relative performance contracts. On a behavioral level, a tournament can filter away the common shock and reduce the magnitude of risk

[4] Both the theoretical and empirical literature on tournaments frequently focuses on two player tournaments (e.g. Lazear and Rosen (1981), Bull, et. al (1987), Schotter and Weigelt (1992), Hvide (2002), among others). While real world tournaments often involve more than two players, the two player case highlights a common criticism of tournaments by agricultural producers; namely that the producers cannot predict what output they must produce to earn the high payoffs.

[5] Risk neutrality of agents is a common assumption in this literature. Indeed, the experimental studies of Bull, et. al (1987) and Schotter and Weigelt (1992) postulated risk neutral agents in their modeling.

faced by agents thereby possibly affecting their behavior. While we have assumed that agents are risk neutral, it might be the case that our experimental subjects may have a variety of risk preferences. By varying the size of our common shock while holding expected payoffs constant, we ought to be able to infer risk aversion by observing changes (if any) in the behavior of our subjects.

Since each agent is identical, they have the same effort cost functions and we assume that the cost functions satisfy the assumptions, $c_i(0) = c(0) = 0$, $c'(e_i) > 0$ and $c''(e_i) > 0$. For our experiment, we will need to impose a specific functional form on the cost functions so we adopt the cost structure used by Bull, et. al. (1987) in their experiment, which is $c(e_i) = \frac{e_i^2}{k}$ where k is a positive constant.

Under the rank order tournament, agent i receives the high payment R if $y_i > y_j$ and the low payment r if $y_i < y_j$.[6] The probability of agent i receiving the high payment is then given by Prob($u_i - u_j > e_j - e_i$) where $u_i - u_j \sim N(0, 2\sigma^2)$ so that the risk agent i faces is twice the idiosyncratic variance. Note that a major benefit of tournaments is that the common shocks cancel out so that each agent only faces idiosyncratic shocks, although this could be better or worse depending how large the variance of the common shock is relative to the total variance. If the common shock variance makes up more than 50% of the total variance, then we have $\sigma_C^2 + \sigma^2 > 2\sigma^2$ so that the agent faces less risk under tournaments since σ_C^2 would be eliminated.

With normally distributed random shocks, Prob($u_i - u_j > e_j - e_i$) =

[6] In the case of a tie, we flipped a coin to determine the winner in our experiments.

1 - $F(e_j - e_i)$ where $F(\bullet)$ is the normal cumulative density function of $u_i - u_j$. Thus, agent *i's* objective function is:

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

which, after some algebra, can be written as:

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

Agent *j's* objective function is:

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

The two agents therefore play a game where their strategies are their effort levels and their payoffs are given by (3) and (4). The first order condition with respect to the effort for each agent is:

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

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

where $f(\bullet)$ is the density function. Conditions (5) and (6) suggest that,

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

so that $e_i = e_j = e^*$ which is a symmetric Nash equilibrium. This also implies that the density function $f(\bullet)$ will be evaluated at zero so that $f(0) = \frac{1}{\sqrt{2\pi(2\sigma^2)}}$, which suggests that the optimal Nash equilibrium effort levels are:

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

For our experiment, we restrict our effort choice set for both agents to be the set of integers from [0, 100] and we chose our parameters carefully to ensure interior solutions in this set.[7]

B. The Fixed Standards Model

For simplicity, we focus attention on a fixed performance standard contract with just binary payoffs so that an agent *i* receives the high payoff *R* if his output exceeds some fixed standard y^* and *r* otherwise. Thus, the probability that agent *i* receives the high payoff is simply Prob($y_i > y^*$), which is equivalent to Prob($u_C + u_i > y^* - e_i$), where $u_C + u_i \sim N(0, \sigma^2_C + u_i)$. Letting $G(\bullet)$ be the cumulative density function 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 under the fixed standards contract is then:

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

with first order condition:

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

Note that since the agent under this contract does not compete against any other player, (10) is just the optimality condition for a standard optimization problem. Because $g(\bullet)$ is a normal density function, solving for the optimal effort level that satisfies (10) will be complicated and will require a numerical solution after we have chosen specific values

[7]Bull, et. al., and Schotter and Weigelt impose the same restriction. We try to maintain consistency with other studies in much of our experimental setup so that we have some basis for comparison when assessing final results. Thus, wherever we can, we choose similar parameters and experimental restrictions to maintain this consistency.

for the parameters.   We will discuss the specific parameters in the next section, but it suffices to say for now that, given the objectives of this study, we selected our parameters to ensure that the agent's ex-ante profits are equal under both tournaments and fixed standard contracts.

III. Experimental Parameters

A. Tournament Parameters

In setting our experimental parameters, we have tried to maintain consistency with prior experimental studies on tournaments. Thus, the specific parameter we choose for our cost function is $c(e_i) = \frac{e_i^2}{k} = \frac{e_i^2}{10,000}$ and $e_i \in [1, 2, ..., 100]$.[8]

Another objective we had was to observe behavior and welfare as we increased the relative size of the variance of the common shock. We hold constant the total variance (sum of the variances for the common and idiosyncratic shocks) at 500 (standard deviation of 22.3) while varying the size of the common shock standard deviation from 0 to 7, from 7 to 15.8, and from 15.8 to 18.7 across different experiments.

In choosing the payments *R* and *r* for the tournament, we had to consider a couple of factors. First, we needed to choose the size of the *spread* between *R* and *r* in order to induce a certain effort level via the incentive compatibility constraint (8). Second, we had to choose the size of the payoff *r* to satisfy some hypothetical participation constraint for the agent. These are also the sorts of constraints that a processor would have to face in designing a contract for growers in practice. To mimic what a real world processor might do, we assume that the processor would like to design a tournament contract to minimize the cost of achieving some performance target or achieving some average level

[8] Each of our experimental subjects faced this same identical cost function.

of performance at minimal cost.[9] The performance target we chose was 37 as it does not appear to be an obvious number that subjects may focus on thereby biasing the results. Moreover, it was the Nash equilibrium effort level chosen by Bull, et. al. so that it facilitates comparison of our results with theirs. Therefore, $R-r$ was chosen to implement an effort level of 37 in order to induce an average output, E(y) =37. Additionally, we assume that r is chosen to ensure that agents' participation constraints are satisfied. While we do not know the actual reservation utilities of our experimental subjects, we did want to ensure an expected payoff of a minimum of \$18.90 per experiment to each of our experimental subjects. Each experiment consisted of four ten round sessions for a total of forty rounds of play. Thus, the per-round participation constraint involves dividing 18.90 by 40 to get .4725.

Now we are ready to derive the optimal values of R and r. With a target effort level of 37, k =10,000, and if the variance of the idiosyncratic shock is half the total variance of 500 (i.e. $\sigma^2 = 250$ or $\sigma = 15.8$), the optimal payment spread is $R - r = .41$. In order to pin down r, we can use the per-round participation constraint,

(11) $$r + f(0)[R-r] - \frac{e_i^2}{k} = r + \frac{1}{2}[.41] - \frac{37^2}{10000} \geq 0.4725$$

[9] Strictly speaking, the principal should choose the target output level as part of an overall profit maximizing strategy. However, given our objective of choosing our experimental design in a way that is relatively consistent with other experimental studies, it was impossible to incorporate profit maximizing levels of effort and payoffs. For example, the effort space is bounded between 0 and 100, but the profit maximizing effort level in a simple principal agent model would yield target effort level of e* = 5000, a payoff spread of $R - r$ = 56.05, and r = 1295. These are well outside the bounds of our experimental constraints. Therefore, we have assumed that the principal is solving some cost minimization problem of achieving a performance target of 37, which is the Nash equilibrium outcome of Bull, et. al's study. While assuming the principal behaves in this way is admittedly arbitrary, we feel that it is not completely out of touch with reality. In practice, performance targets are often established by long run investments in fixed factors, such is processing capacity, and/or by product or branding choices that require minimal quality standards in inputs. Moreover, contract design is often a short run problem so that processors may not actually choose simultaneously both the performance target and the cost minimizing contract parameters for achieving that target. Thus, we believe that our approach of assuming that the principal minimizes contracting costs conditional on some established performance target is not unreasonable for dealing with short run welfare issues.

Solving for r yields $r \approx .40$.[10] This implies that, in each round, each agent could choose an effort integer from 0,…,100 and to this effort level is added an idiosyncratic shock, u_i, and an aggregate shock, u_C, which are distributed $u_i \sim N(0, 250)$ and $u_C \sim N(0, 250)$, to get $y_i = e_i + u_C + u_i$. The output for agent j is similarly defined. If $y_i > y_j$, then agent i gets $R = .81$ and agent j gets $r = .40$, and if $y_i < y_j$, then agent i gets $r = .40$ and agent j gets .81. 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.[11]

We also calibrated payoffs for a case when the common shock is large (70 percent of total variance) still holding total variance at 500. In this case, $\sigma = 18.7$ and the pay spread that would implement an effort level of 37 would be $R - r = .32$. The reservation utility for the participation constraint is once again 18.90 or 0.4725 per round,[12] and solving for r yields $r = .45$. We also approximated two normal distributions, one for the common shock with a standard deviation of 18.7 and another for the idiosyncratic shock of 12.2 using the same method we used earlier.

[10] We say "approximately" 40 because our numerical calculations had minor rounding errors. For example, effort was actually 36.83 for an idiosyncratic variance of 250 and a pay spread of .41. The expected payoffs were also slightly different from .4725 due to minor approximation errors but the payoff did not deviate by more than 0.001.

[11] 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.

[12] There were minor rounding errors here so the expected payoff for the agent was slightly above .4725 per round. But the deviation was only about 0.0016.

A third experiment reduced the variance of the common shock to only 10% of the total variance of 500 so that the common shock standard deviation is only 7.07. The pay spread that would implement 37 is $R - r = .55$ and the level or *r* that would satisfy the participation constraint with reservation utility set at .4725 is $r = .33$. We approximated two normal distributions, one for the common shock with a standard deviation of 7.07 and another for the idiosyncratic shock using a standard deviation of 21.2.

A fourth experiment eliminated the variance of the common shock altogether so that total variance is strictly due to the idiosyncratic shock. This experiment would replicate (with minor changes in assumptions distributional assumptions) Bull, et. al's study mentioned earlier.

The four sets of tournament parameters and payoffs generated above allow us to assess subject behavior under varying sizes of the common shock. Since we have adjusted payments appropriately to ensure that the Nash equilibrium effort level is 37 and that expected payoffs are identical at approximately .4725 across the four sizes of the common shock, we should theoretically observe no differences in behavior under risk neutrality. This provides us with our first hypothesis:

> *Hypothesis 1: Changes in the size of the variance of the common shock (as a fraction of total variance of 500) should not alter equilibrium behavior under risk neutrality once compensating payment spreads R – r are made.*

B. Fixed Performance Parameters

In calibrating parameters for the fixed performance contract, we had two goals in mind. First, we wanted the optimal effort level to be identical to the Nash equilibrium effort of 37 under tournaments. This allows us to remain consistent with our assumption that the principal is interested in implementing this effort level to achieve a target or expected output level of 37. Second, we assumed that if the principal were forced to use a fixed standards contract due to say, a ban on tournaments, it would still have to satisfy the same participation constraints for agents so that we maintain the reservation utility of .4725 per round in calibrating fixed performance payments.

Note that the incentive compatibility constraint for the agent is given by (10) so that given a choice of $e_i = 37$, we can solve for the payment spread.[13] However, before doing so we need to choose a fixed standard y^* which output must exceed in order for the agent to receive the high payment R. If output falls below y^*, then the agent receives r. An obvious choice would be $y^* = 37$, but we avoided this choice because we did not want to provide our experimental subjects with a focal point so that they naturally gravitate toward the optimal solution, 37. Instead, we chose $y^* = 41$ and then adjusted our payment spread so as to ensure that 37 is the optimal effort level. Since $y^* > 37$, $g(\bullet)$ in (10) does not simplify into an easily manageable form as in the tournament model. We therefore numerically solved for the optimal wage spread which is $R - r = .42$. The value of r that would result in an expected payoff of .4725 per round to satisfy the participation constraint is $r = .43$. This leads to our next hypothesis.

Hypothesis 2: Under the parameters chosen, effort level under the tournament contract and the fixed standard contract should be identical, on average.

Of course, this hypothesis relies on the assumption that agents are risk neutral.

Since the fixed standard does not filter away the common shock, the total risk faced by the agent is $\sigma_C^2 + \sigma^2$. However, unlike the tournament case, varying the size of

[13] We also evaluated the second order conditions at 37 to ensure that we are at a maximum.

the common shock (as a relative proportion of the total shock) should have no impact on agents' risk total risk exposure, since we hold total variance constant at 500. This leads to our third hypothesis:

> *Hypothesis 3: Increasing the size of the variance of the common shock (as a fraction of total shock) should not alter agents' behavior under a fixed performance contract.*

C. Agent Welfare

Under the parameters in the previous sections, ex ante agent welfare was established to be .4725 dollars per round. Calculating ex post welfare for both tournaments and fixed performance was fairly straightforward as it simply involved taking the amount earned in each round (R or r) and subtracting the cost of effort from the amount earned. Since the ex-ante welfare was set to be .4725 across both the tournament and the fixed standard contract, our welfare hypothesis for the agents is:

> *Hypothesis 4: The ex-post welfare under the tournament should not differ significantly from the ex-post welfare under the fixed standard contract.*

D. Principal Welfare

On the cost side, it is straightforward to see that one advantage of the tournament for the principal is that there is no uncertainty with regard to total payments that must be made to two agents. That is, in a two agent tournament, the cost to the principal is always $R + r$ since there is always one winner and one loser so there is little uncertainty with regard to total payments made to the two agents. On the other hand, if the principal must contract with the same two agents but must use a fixed standards contract, the expected total payments can be (R, R), (r, r) or (R, r). The expected total payment would be:

(12)

$$E(pay) = \left(1-G(y^*-e^*)\right)^2(2R) + 2\left(1-G(y^*-e^*)\right)G(y^*-e^*)\left(R+r\right) + G(y^*-e^*)^2\left(2r\right)$$

Consequently, one economic rationale for explaining the prevalence of tournament contracts in certain agricultural sectors is that it provides processors with a mechanism for reducing the variance in its costs of procurement.[14]

While minimum cost variation is a desirable goal, it is unlikely that principals care only about this per se; that is, they also care about profits and profit variation. To generate insights about processor profits, we had to make some heuristic assumptions about the shape of the revenue function since we have little information about how output may relate to revenues in practice. As mentioned in Section III A, we did not solve for the long run profit maximizing performance target and instead assumed that, in the short run, processors try to minimize the cost of achieving some pre-set performance target or average level of output.[15] Thus, we can interpret the revenues in this section to be short run revenues (e.g. flock to flock revenues).

To construct processor payoffs, we heuristically assumed two types of possible short run revenue functions.[16] This allows us to determine whether qualitative comparisons of ex-post payoffs between tournament and fixed standard incentives will remain robust under different revenue structures. The first revenue function is just the

[14] This would require that at least some processors be risk averse.

[15] For example, broiler companies are unlikely to change their processing capacity or quality performance standards on a per flock basis. Instead, quality targets are determined by fixed investments in capacity or product line over the long run. In this case, contracts will be designed to minimize the costs of inducing growers to achieve these performance targets and short run (flock by flock) profits will be based on some function of output minus payments made to growers.

[16] While we will not defend our quantitative results on the basis of the rigor of our process, our analysis can, nonetheless, provide general, qualitative insights about whether processor welfare might be higher under tournaments or fixed standard contracts, and whether the qualitative results would remain robust under alternative revenue functions.

simple revenue function commonly found in the agency literature which is that output is equivalent to revenue so that $R(y) = y$. However, because the actual unit of measure for y is not given in dollar terms in our experiment, it is difficult to interpret output as revenue. Instead, we make short run revenue a linear function of y by scaling y down appropriately so as to facilitate a reasonable comparison between revenues and cost. The scaling factor we choose was 0.1 so that expected per round revenue would just be $R(y) = 0.1y$ so that the expected ex-ante revenue would be \$3.7 per round. Also, with two agents, we assume that the processor cares about the average output so that we actually use $R(y) = 0.1\overline{y}$ where $\overline{y} = \frac{y_1 + y_2}{2}$. As a result, the ex post payoff for the processor under the tournament would just be:

(13) $$\pi_1^t = .01\overline{y} - (R + r)$$

For the fixed performance contract, the payoff would be:

(14) $$\pi_1^f = .01\overline{y} - (p_1 + p_2)$$

where $p_i \in \{R, r\}$ for i =1, 2. We should note that, while total payments made by the principal are deterministic under the tournament and random under the fixed standard contract, it may be the case that even a risk averse processor may prefer the fixed standard contract because the randomness of the total payments offers a natural hedge against random revenues. This is because when average output is low and thereby revenue is low, there is also a higher probability that payments will be low. Thus, low revenues are partly compensated by low payments.

The second revenue function we look at is a “loss” or “distance” function of the form $R(y) = R(\hat{y}) - \delta\left|\hat{y} - y\right|$ where $\hat{y}$ is some minimal performance threshold that output must exceed or there will be a loss in the amount $\delta|\hat{y}-y|$ which is subtracted from the revenue, $R(\hat{y})$. The revenue $R(\hat{y})$ then represents the maximum achievable revenue when performance objectives are met. Loss functions have been used in some of the recent contracting literature (e.g. Dessein, 2002), and can be justified from a practical standpoint in that processors often attempt to meet some minimal performance standard rather than push for maximum performance. For example, plant capacity often dictates that some minimal flow of output must be available within any given time period to reduce efficiency losses. Once the supply of output has met the minimal flow requirements, having additional output does not add much more to revenue. Similarly, quality consistency in food products often requires that the raw input commodity meet certain minimal quality standards. But pushing for quality beyond minimal standards would increase costs but not add much to revenue.

The ex post payoff with our second revenue function for a tournament contract is,

(15) $$\pi_2^t = R(\hat{y}) - \delta|\hat{y}-\bar{y}| - (R+r)$$

We chose $\delta = 0.1$ and $R(\hat{y}) = 3.7$ to be consistent with our first revenue function. For simplicity, we assume that $\hat{y} = 37$ so that some loss occurs whenever $\bar{y}$ falls below 37. The ex post profits using the tournament is calculated using,

(16) $$\pi_2^t = 3.7 - 0.1|37-\bar{y}| - (R+r)$$

For the fixed standard contract, it is calculated using,

(17) $$\pi_2^f = 3.7 - 0.1|37-\bar{y}| - (p_1 + p_2)$$

where $p_i \in \{R, r\}$ for i =1, 2.

These models allow us to conduct an empirical exercise to understand qualitative issues such as (1) whether ex-post principal profits are greater under tournaments or fixed standard contracts, and (2) whether the qualitative results would be robust to alterations in the revenue functions. This will shed light on why processors prefer tournaments in certain industries but not in others and how a ban may affect short run profits of processors.

IV. The Experiments

To test our hypotheses, we ran a total four experiments on four separate days. Each experiment used a different set of subjects and a different size common shock variance. Each experiment involved 12 students[17] who were recruited at The Ohio State University via posters and/or email lists across several departments.[18]

For each experiment, subjects would arrive in a room where we randomly assigned each subject to one of 12 chairs until all chairs were occupied. We informed the subjects that they had an opportunity to earn money and the amount they earned was dependent on the decisions they made during the course of the experiment. The first session of the night involved a tournament experiment, where each subject was paired with another subject, although the identities of the pair members were not revealed to the parties. The subjects would then play 10 identical rounds of this tournament game, where in each round, each subject was asked to choose a “decision number” (effort) from 0 to 100. The higher the number a subject chose, the higher the cost of that decision to the

[17] Four of the subjects participated in two of the experiments. But the other 44 subjects participated only in one of the experiments.
[18] Including the agricultural economics department, the business school, and several social and physical science departments.

subject.[19] After the decision numbers were chosen the subject would enter these numbers into their worksheets and an administrator would record the decisions in a computer. Subsequently, one subject would draw the "common shock" number from a bucket with frequencies that approximated a normal distribution[20] and all subjects in the room were asked to add this number to their decision. Then each subject would individually draw a random number from another bucket with frequencies approximating a normal distribution, and then these individual numbers were also added to their decision plus the common shock. The sum of the decision number, the common shock, and the idiosyncratic shock would be their "output". For each matched pair of subjects, the administrator would compare outputs and the pair member with the higher output would receive the high payment *R* while the other pair member gets *r*. Each subject would record his/her winnings in the worksheet and subtract their decision cost to get his/her net earnings from that round. After the round ended, the next round began and the entire process was repeated. There were a total of ten rounds and all rounds were identical. At the end of the tenth round, the subjects calculated their payoffs for the ten round session where the total payoff was just the sum of the net payoffs per round. All subjects received the same cost sheet, knew the distribution of the numbers in the buckets, and all other experimental parameters. Only the identity of the pair members was not common knowledge. A session typically last between 20-25 minutes. Example instructions to the tournament are contained in Appendix A.

---

[19] Each subject was given a cost table where the costs were calculated using the cost function $c(e_i)=\frac{e_i^2}{10{,}000}$ specified in Section IIIA.

[20] We described in Section IIIA how we approximated the normal distributions.

It should be noted that, while the tournament was repeated over 10 rounds, the theory is based on a static model. Nonetheless, such repetition is common in experimental practice since subjects make complex decisions. Moreover, the only subgame perfect Nash equilibrium to a finitely repeated game involves the choice of the Nash equilibrium decision levels to the one-shot game in each round. Thus, predictions concerning equilibrium play was independent of finite repetition (Bull, et. al.).

Once the ten round tournament session was completed, we started a second experimental session for the fixed standard contract. In all respects, this session was identical to the tournament session, except that each subject played against a fixed standard of $y^* = 41$, rather than against a pair member. Example instructions are contained in Appendix B.

Once the first two sessions were completed, we conducted another two sessions, a tournament session and a fixed standards session. However, the subjects didn't 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 ten highest bidders for the second tournament session got to participate in the post-auction tournament. Similarly, the ten highest bidders for the fixed standard session got to participate in the additional session. Example instructions and description of the auction is in Appendix C.

We conducted the second half sessions and the auction for two reasons. First, Friedman and Sunder (pages 98 and 99) point out that learning effects are pervasive in experiments so that we wanted to have subjects repeat the sessions at least once so that some of the learning effects are neutralized in the second half sessions. Secondly, some subjects were not very motivated and did not seem to make decisions very carefully, which may contaminate the overall results. By having all subjects bid into the second half sessions, we were able to select out some of the less motivated subjects who revealed their low interest in playing by bidding a low amount to continue. Consequently, we expected the second half sessions (one tournament and one fixed standard) to produce more consistent data with fewer outliers caused by learning and unmotivated subjects.

V. Results

A. Effort Levels

Table 1 provide experimental results for the decision numbers (effort) chosen by the experimental subjects across four experiments. The data is segmented into a pre-auction and post auction sessions for both tournaments and fixed performance. Additionally, for the pre-auction sessions, separate summary statistics were provided for rounds 1-5 and rounds 6-10. We do not partition the post-auction data in this manner since learning effects for early rounds should be less of an issue; all subjects have already played in an identical pre-auction round.

Using data from all four experiments (row 1), the average effort level was close to the Nash equilibrium of 37 for all three partitions of the data.[21] For the pre-auction data, the average decision levels were 35.8 and 35.4 for the 1-5 and 6-10 rounds, respectively. For the post-auction data, the mean was a remarkable 37.1 which is almost exactly the Nash equilibrium level predicted in our theoretical model. The associated standard deviations for rounds 1-5 and 6-10 are 19.8 and 18.6, respectively. The post-auction standard deviation was 16.2. Based on casual observation of summary statistics, it appears that the subjects come close, on average, to the predicted optimal strategy of 37

[21] The means were calculated across players and the specified rounds.

and that they appear to improve their play (closer to Nash and reduced standard deviation) in the post-auction sessions where learning effects are not as pronounced.

Table 1: Summary Statistics for Effort Levels

	Pre-Auction Session Means		Post-Auction Session Means	Pre-Auction Session Standard Deviations		Post-Auction Session Standard Deviations
Data	Rounds 1-5	Rounds 6-10	All Rounds	Rounds 1-5	Rounds 6-10	All Rounds
1.Overall Tournament	35.8	35.4	37.1	19.8	18.6	16.2
2.Tournament with $\sigma_c = 0$	42.9	38.6	39.7	22.8	17.2	18.6
3.Tournament with $\sigma_c = 7$	35	36.8	40.1	19.3	20.6	18.5
4.Tournament with $\sigma_c = 15.8$	33.2	36.5	36.6	17.8	18	12.4
5.Tournament with $\sigma_c = 18.7$	32.2	29.7	32.1	17.5	17.6	13.0
6.Overall Fixed Std.	42.3	39.6	40.1	15.7	16.3	14.1
7.Tournament with $\sigma_c = 0$	37.4	35.4	36.4	16.3	20.4	14.7
8.Tournament with $\sigma_c = 7$	45.5	40.4	43.3	17.7	17.7	15.5
9.Tournament with $\sigma_c = 15.8$	45.7	41.9	40.6	11.6	12.7	11.5
10.Tournament with $\sigma_c = 18.7$	40.5	40.9	39.9	15.3	12.8	13.7

Table 1 also provides summary statistics for the fixed standard sessions. Using data from all four experiments (row 6), the average effort level was slightly above the optimal effort choice of 37 for all three partitions of the data. For the pre-auction data, the average decision levels were 42.3 and 39.6 for the 1-5 and 6-10 rounds, respectively. For the post-auction data, the mean was 40.1 which is slightly higher than the optimal choice of 37. The associated standard deviations for the pre-auction rounds 1-5 and 6-10 were 15.7 and 16.3, respectively. The post-auction standard deviation was 14.1.

Hypothesis 2 states that there should be no significant difference between the effort levels chosen under the tournament and effort levels chosen under the fixed standard contract. The summary statistics from Table 1 suggests that effort level was slightly higher under the fixed standard contract than under the tournaments. To test this theory, Wilcoxon tests were conducted separately for the pre-auction rounds 1-5 and rounds 6-10, and for the post-auction data. The p-values were 0.00 for all three partitions of the data so that we would reject Hypothesis 2 on the basis of this test.

Because our subjects repeated rounds and sessions, we have numerous repeated observations on each experimental player. Moreover, some of our subjects in our fourth experiment (with $\sigma_c = 0$) had already participated in an earlier experiment. One downside to this is that many of the across round observations will not be a part of a random sample. However, the repeated observations do allow us to use fixed effects to control for unobservable heterogeneity, such as rate of learning, risk tolerances, etc, which may affect the way subjects choose their strategies. We run three different regressions with subject fixed effects to control for unobservable heterogeneity, with each regression corresponding to a different partition of the data. The dependent variable for each regression is effort, and the explanatory variables are a tournament dummy, which equals 1 whenever the data was generated from a tournament session and zero otherwise; a variable for the standard deviation of the common shock which can take on four values, 0, 7, 15.8, and 18.7; a post auction dummy, which equals 1 if the data came from a post auction session; and interactions terms for the tournament dummy and common shock, the post auction dummy and the common shock, and the tournament dummy and post auction dummy. Finally, all regressions contain 43 fixed effects since there were a total of 44 subjects across all four experiments (4 subjects were repeat participants).

Table 2 contains the results of the regression using three different partitions of the data. Regression 1 uses all experimental data. Regression 2 uses all data except the first five rounds of the pre-auction data to reduce the impact of learning effects. Regression 3 uses only post-auction data to reduce learning effects and subject outlier effects. Since these regressions control for unobserved heterogeneity across subjects, they can be used to conduct a more sophisticated assessment of Hypothesis 2.

Regression (1), which uses all the data, yielded a tournament coefficient of 0.87 but the p-value was 0.54 so that, by itself, the tournament dummy appears to have no impact on effort. However, the interaction term between the tournament dummy and the common shock standard deviation is -0.59 and with a p-value of 0.00. This suggests that, under tournaments, effort levels are decreasing in the standard deviation of the common shock. This result was fairly robust to different partitions of the data as no qualitative changes occurred in regressions (2) and (3). This suggests that Hypothesis 2 can be rejected for non-zero values of common shock standard deviation.

Recall that Hypothesis 1 and Hypothesis 3 were statements about the way changes in the relative size of the common shock variance may or may not affect decisions of subjects. The null hypotheses would be that adjusting the relative size of the variance of the common shock will not alter behavior either under the tournament contract or the fixed standard contracts after payments are adjusted to compensate for differences in the size of the common shock variance. Essentially, these hypotheses test for the behavioral relevance of varying the size of the common shock. We anticipate that behavior under the tournament contract is more likely to be affected by changes in the size of the common shock since such changes alter the risk faced by subjects under tournaments. This is because the tournament filters out the common shock so the larger the common shock, the more risk is filtered out by the tournament. On the other hand, altering the relative size of the common shock, while holding the variance of the total shock fixed, will have no impact on the amount of risk faced by subjects under the fixed standard contract.

For each of the four experiments we performed, we altered the relative size of the standard deviation of the common shock in our normal distribution approximations for our experimental random draws. Then we made adjustments to the spread between high payments, *R*, and low payments, *r*, to ensure that 37 was the Nash equilibrium. This provides us with the data to refute hypotheses 1 and 3. Table 1 provides the summary statistics for different sizes of the standard deviations of the common shock. Under the tournaments, the average effort levels do appear to be responsive to changes in the size of the common shock standard deviation. Regardless of the data partition, average effort levels appear to be declining in σ_c. For example, for rounds 1-5 in the pre-auction experiments, higher average effort levels appear to be associated with lower common shock standard deviations. For $\sigma_c = 0$, mean effort level was 42.9, while for $\sigma_c = 19$, it dropped down to 32.2. Similar patterns were observed for rounds 6-10, and for the post auction sessions as can be seen from Table 1, rows 2-5. To test our hypothesis formally, we conducted a Kruskal-Wallis test for equality of means across several samples.[22] For the pre-auction, rounds 1-5 data, the KW test yielded a Chi-square(3) test statistic of 6.79

[22] This is a multiple sample generalization of the Wilcoxon test.

with an associated p-value of 0.08. Thus, the null hypothesis that the mean effort level should be equal across alternative sizes of the common shock standard deviation is rejected at the 10% level but not rejected at the 5% level.[23] For rounds 6-10, the Chi-square(3) test statistic was 7.21 with a p-value of 0.07. Thus, our conclusions do not change by using only the later rounds data. However, for the post auction tournament sessions, the same test yielded a Chi-square (3) test statistic of 20.51 so that the null could be rejected even at the 1% level. Thus, we reject our Hypothesis 1.

Turning our attention toward the fixed standard contract, we can see from Table 1, rows 7-10 that average effort level did not exhibit the declining pattern it did under tournaments for increasing sizes of the common shock standard deviation. While there was still substantial variation in mean effort levels, it may have had more to due with sampling error across different subjects rather than variations in the common shock standard deviation. The Kruskal-Wallis on the pre-auction, rounds 1-5 data yielded a Chi-square(3) test statistic of 12.03 with an associated p-value of 0.00 so we can reject our null of equality of effort across common shock variances. For the rounds 6-10 data, the Chi-square(3) test statistic was 3.18 with an associated p-value of 0.36 so that we could not reject the null of equal effort levels. Finally, for the post-auction fixed standards sessions, the Chi-square(3) was 17.73 with a p-value of 0.00 so that we reject the null of equal effort levels. One can see that the results are more mixed for fixed standard contracts than tournaments contracts.

Turning our attention to the regression results in Table 2, we find that the estimate of the coefficient for the standard deviation of the common shock was 0.67 with a p-value

[23] It should be noted, however, that in these early rounds, learning effects may be substantial so that our sample across rounds may not be efficient thereby possibly biasing our results toward not rejecting the null.

Table 2: Effort Regression Results (Effort is dependent variable)

	Data Used		
Variables	(1) Using all data	(2) All data except pre-auction rounds 1-5	(3) Post auction data only
Tournament dummy (1 if data from a tournament session, 0 otherwise)	0.87 (0.54)	1.72 (0.33)	1.69 (0.18)
Common Shock Std Deviation	0.67 (0.00)	0.53 (0.02)	0.01 (0.96)
Post Auction Dummy (1 if data from a post-auction session)	-0.12 (0.93)	1.82 (0.32)	--
Tournament × Common Shock	-0.59 (0.00)	-0.57 (0.00)	-0.60 (0.00)
Post Auction × Common Shock	-0.05 (0.59)	-0.10 (0.37)	--
Tournament × Post Auction	2.15 (0.08)	0.79 (0.59)	--
Fixed Effects for experimental subjects.	Too many to list (44 total subjects)	Too many to list (44 total subjects)	Too many to list (38 total subjects)
No. of Observations	1760	1280	800
R-squared	0.406	0.478	0.62

Note 1: All regressions were estimated using the robust Huber-White sandwich estimator.
Note 2: p-values are contained in the parentheses below the coefficients.

of 0 in regression 1. Since it is positive and significantly different from zero, it appears that larger common shock standard deviations are associated with higher effort. However, for the post-auction data only, this same coefficient is not significantly different from zero. Since the latter is the more reliable data in our estimation, we give preference to these results. The interaction term between common shock standard deviation and the tournament dummy in the post-auction data is negative and highly significant. What we can interpret from regression 3 is that when the tournament dummy is 0 (fixed standard contract), then the coefficient on the common shock standard deviation is insignificantly different from zero so that Hypothesis 3 cannot be rejected. That is, increasing the relative size of the common shock will have no impact on effort under the fixed standard contract. However, when the tournament dummy is 1 (tournament contract), the coefficient on the common shock standard deviation is -0.60 and it is significantly different from zero. Hence, we would reject Hypothesis 1. Indeed, this confirms our earlier conjecture that effort is declining in the common shock standard deviation under tournaments.

B. Agent Welfare

A crucial aspect of this study is that we attempt to gain insights about how the welfare of agents might be affected by tournaments and fixed standard contracts. This will allow us to make inferences about how growers might be impacted by alternative contractual regimes and shed light on their attitudes toward these contracts. Recall from Section III.C that we had set agent reservation utility at 0.4725 dollars per round so that this represents our expected payoffs for our subjects. Hypothesis 4 states that ex post welfare should be equal to 0.4725 under both the fixed standard and tournament contracts.

Table 3 contains summary statistics on the actual per-round earnings of subjects in our four experiments. A casual glance of the data reveals that, for the most part, mean per-round earnings were higher and standard deviations were lower under fixed standard contracts. To formally test Hypothesis 4, we also conducted Wilcoxon tests for the overall data, the pre-auction rounds 1-5 and rounds 6-10, and for the post-auction data. The p-value for the overall data was 0.00 which rejects the null hypothesis that per-round net earnings are equal under tournaments and fixed standard contracts. For the pre-auction rounds 1-5 data, the p-value was 0.03 which also rejects the null. However, for rounds 6-10, the p-value was 0.33. Finally, using only the post-auction data yielded a p-value of 0.09, so that the null can only be rejected at the 10% level.

Table 3: Summary statistics for per-round earnings by subjects (agents).

Data	Pre-Auction Session Mean Earnings		Post-Auction Session Mean Earnings	Pre-Auction Session Standard Deviations of Earnings		Post-Auction Session Standard Deviations of Earnings
	Rounds 1-5	Rounds 6-10	All Rounds	Rounds 1-5	Rounds 6-10	All Rounds
1.Overall Tournament	0.439	0.446	0.442	0.237	0.238	0.247
2.Tournament with $\sigma_c = 0$	0.37	0.427	0.413	0.315	0.313	0.304
3.Tournament with $\sigma_c = 7$	0.446	0.428	0.41	0.262	0.268	0.289
4.Tournament with $\sigma_c = 15.8$	0.463	0.44	0.456	0.176	0.178	0.204
5.Tournament with $\sigma_c = 18.7$	0.476	0.492	0.49	0.15	0.157	0.153
6.Overall Fixed Std.	0.486	0.469	0.479	0.198	0.191	0.195
7.Tournament with $\sigma_c = 0$	0.439	0.453	0.494	0.188	0.187	0.187
8.Tournament with $\sigma_c = 7$	0.444	0.481	0.50	0.204	0.183	0.185
9.Tournament with $\sigma_c = 15.8$	0.488	0.449	0.47	0.197	0.21	0.192
10.Tournament with $\sigma_c = 18.7$	0.572	0.492	0.45	0.175	0.185	0.213

A more sophisticated test involves looking at the regression coefficients in Table 4 since these results controls for unobserved heterogeneity across subjects. Regression (1), which uses all the data, yielded a tournament coefficient of -0.07 with a p-value of 0.003 so that we would reject the null that the tournament and fixed standards contracts should result in the same level of net earnings. However, the coefficient for the interaction term between the tournament dummy and the common shock standard deviation variable is positive and significant. Thus, it appears that tournaments have a negative impact on per round net pay but this negative impact would be partially offset by increasing the size of the common shock standard deviation. Intuitively this makes sense since a relatively higher standard deviation implies that more risk gets filtered out by tournaments so that

Table 4: Agent Net pay Per-Round Regression Results.

	Data Used		
Variables	(1) Using all data	(2) All data except pre-auction rounds 1-5	(3) Post auction data only
Tournament dummy (1 if data from a tournament session, 0 otherwise)	-0.07 (0.003)	-0.08 (0.01)	-0.11 (0.00)
Common Shock Std Deviation	0.000 (0.99)	-0.004 (0.29)	-0.01 (0.10)
Post Auction Dummy (1 if data from a post-auction session)	0.04 (0.07)	0.02 (0.39)	--
Tournament × Common Shock	0.003 (0.03)	0.005 (0.00)	0.006 (0.00)
Post Auction × Common Shock	-0.003 (0.02)	-0.001 (0.61)	--
Tournament × Post Auction	0.00 (0.99)	-0.01 (0.59)	--
Fixed Effects for experimental subjects.	Too many to list (44 total subjects)	Too many to list (44 total subjects)	Too many to list (38 total subjects)
No. of Observations	1760	1280	800
R-squared	0.08	0.07	0.09

Note 1: All regressions were estimated using the robust Huber-White sandwich estimator.
Note 2: p-values are contained in the parentheses below the coefficients.

agents face less uncertainty and may be able to make more efficient decisions. These qualitative results remained robust under regressions (2) and (3), which used smaller subsets of the data.

Based on our statistical results, we would conclude that agents generally earned less profit under tournaments relative to fixed performance contracts but this result is mitigated as the size of the common shock increased. This is consistent with the tournaments literature which often cites the presence of large common shocks as a major advantage to using relative performance contracts. We would therefore infer from these results that growers may generally be worst off under tournament contracts, all else being equal, but if common shocks are large, then tournaments may be less harmful and may in some cases even be beneficial to grower welfare. For example, if we revisit the data in

Table 3, we can see that mean net earnings are higher and standard deviations of net earnings are lower under tournaments when the common shock is large at 18.7.

Overall, our results provide a rational explanation for why many growers dislike tournament contracts; simply put, they may earn less money. However, tournaments may be advantages in situations where the size of the common shock variance overwhelms the idiosyncratic variance.

C. Principal Welfare

In Section III D, we outlined a strategy for heuristically modeling short run profits for principals using two different revenue functional forms. These profits are based on teams of two agents with the revenues a function of the average output of the two agents. Table 5 provides summary statistics of per-round earnings for the principal under the revenue/profit function consistent with (13) and (14), where revenue is just a linear function of output. One can easily see that processor revenues are higher under the fixed standard contract in almost every partition of the data. Moreover, the standard deviations of profits also tend to be lower under fixed standard contracts in the majority of partitions.

Indeed, Wilcoxon tests for the overall pre-auction rounds 1-5, the pre-auction rounds 6-10, and the post auction rounds yielded p-values of 0.00, 0.009, and 0.045, respectively. Thus, we conclude that per-rounds profits are not equal under tournaments and fixed performance contracts using the first type of revenue function. Therefore, this data provides us with little empirical evidence for explaining why tournament contracts are used by processors in certain industries.

Table 5: Summary Statistics for per-round earnings for Principals – Revenue Function 1

Data	Pre-Auction Session Mean Earnings Rounds 1-5	Pre-Auction Session Mean Earnings Rounds 6-10	Post-Auction Session Mean Earnings All Rounds	Pre-Auction Session Standard Deviations of Earnings Rounds 1-5	Pre-Auction Session Standard Deviations of Earnings Rounds 6-10	Post-Auction Session Standard Deviations of Earnings All Rounds
1.Overall Tournament	1.98	2.14	2.45	2.45	2.31	2.16
2.Tournament with $\sigma_c = 0$	3.32	2.71	2.58	3.39	2.77	2.57
3.Tournament with $\sigma_c = 7$	1.57	2.04	2.30	1.85	2.29	2.33
4.Tournament with $\sigma_c = 15.8$	1.33	2.25	2.72	1.92	2.22	1.82
5.Tournament with $\sigma_c = 18.7$	1.69	1.54	2.18	1.88	1.83	1.86
6.Overall Fixed Std.	3.17	2.86	2.88	2.00	1.81	1.91
7.Tournament with $\sigma_c = 0$	2.13	2.28	2.59	2.06	1.98	1.54
8.Tournament with $\sigma_c = 7$	3.16	2.99	3.25	2.09	1.70	2.02
9.Tournament with $\sigma_c = 15.8$	3.85	2.84	2.98	2.13	1.66	1.62
10.Tournament with $\sigma_c = 18.7$	3.55	3.32	2.62	1.21	1.83	2.28

Table 6 presents summary statistics for short run principal profits under our second revenue/profit function which is specified in (16) and (17). Here, the results are more mixed as profits are greater under tournaments for some partitions of the data. Specifically, for $\sigma_c = 15.8$ and $\sigma_c = 18.7$, profits are slightly greater under tournaments although standard deviation of profits are still lower under fixed standard contracts. Because of the mixed results, we conducted more detailed Wilcoxon tests than we did under the first revenue function. Specifically, we test for the equality of profits under tournaments and fixed standards for the overall data as well as for the different sizes of common shock standard deviations.

Table 6: Summary Statistics for per-round earnings for Principals – Revenue Function 2

Data	Pre-Auction Session Mean Earnings Rounds 1-5	Pre-Auction Session Mean Earnings Rounds 6-10	Post-Auction Session Mean Earnings All Rounds	Pre-Auction Session Standard Deviations of Earnings Rounds 1-5	Pre-Auction Session Standard Deviations of Earnings Rounds 6-10	Post-Auction Session Standard Deviations of Earnings All Rounds
1.Overall Tournament	1.23	1.38	1.59	1.28	1.43	1.28
2.Tournament with $\sigma_c = 0$	1.53	1.48	1.49	1.12	1.33	1.46
3.Tournament with $\sigma_c = 7$	1.19	1.36	1.46	1.27	1.57	1.47
4.Tournament with $\sigma_c = 15.8$	0.99	1.50	1.86	1.47	1.43	1.02
5.Tournament with $\sigma_c = 18.7$	1.21	1.19	1.57	1.25	1.41	1.11
6.Overall Fixed Std.	1.77	1.82	1.69	0.89	0.79	0.89
7.Tournament with $\sigma_c = 0$	1.39	1.40	1.77	1.33	1.11	0.84
8.Tournament with $\sigma_c = 7$	1.75	1.94	1.71	0.69	0.63	0.89
9.Tournament with $\sigma_c = 15.8$	1.84	1.88	1.84	0.82	0.77	0.74
10.Tournament with $\sigma_c = 18.7$	2.11	2.06	1.47	0.27	0.32	1.02

Using all of the post-auction data, the Wilcoxon test delivered a p-value of 0.00 which suggests that profits are not equal under tournaments and fixed standard contracts. For common shock standard deviations sizes of 0, 7, 15.8 and 18.7, the Wilcoxon p-values were 0.05, 0.11, 0.02, and 0.10, respectively. These results are somewhat mixed at the 5% level but at the 10% level, we can state more confidently that profits are not equal under the two contract types.

An interesting facet of these results is that profits for tournaments appear to be higher as the size of the common shock standard deviations get larger. For example, average profits are higher under the fixed standard contract for $\sigma_c = 0$ and $\sigma_c = 7$, but higher under the tournament when $\sigma_c = 15.8$ and $\sigma_c = 18.7$. These results are consistent with what we might expect which is that the benefits of a tournament are greater as the size of the common shock gets larger.

A possible implication of these results is that different types of operations may favor different types of contracts. For example, for firms that have revenue functions that are close to linear in output where more is always better, the fixed standard contract may be preferred. However, for a firm with a nonlinear revenue function, such as a loss function, more is not always better; indeed the firm only needs to meet some threshold. Once some minimal performance standard is met, then greater output may not add anymore to revenues. We can see that these two types of revenue functions can yield possibly different qualitative results, which may yield insights into why tournaments are prevalent in some industries but not in others.

VI. Conclusion

The purpose of this research was to conduct an experimental analysis of behavior under tournaments and fixed performance contracts. The insights derived from such experiments can shed light on recent controversies surrounding legislative proposals to ban tournament contracts. Specifically, we were interested in understanding how a ban on tournaments might affect ex-post efficiency and welfare of principals and agents.

With regard to issues of efficiency, our statistical tests showed that effort level is significantly lower under tournaments contracts and the discrepancy increased as the size of the common shock standard deviation increased. Thus, even though there should have been no ex-ante difference in effort levels under the two types of contracts (based on our choice of parameters), it appeared that, ex-post, the fixed standard contract was more efficient in that it induced a higher average level of effort from experimental subjects. A particularly interesting facet of the study is that effort is declining with the size of the common shock under tournament contracts. While this may seem counterintuitive, it can be explained by simple economic logic. Essentially, when common shocks variances increase, tournaments tend to be more effective at reducing agents' risk exposure. When risk is reduced, risk averse agents will tend to decrease their effort levels to insure against bad shocks which may cause their performance to drop. While our experimental parameters were generated using risk neutral models, we do not rule out the possibility that our actual subjects may behave in a risk averse manner. We intend to investigate this possibility further in our future research and attempt to infer subjects' coefficients of risk aversion from our experimental data.

Agent welfare (net pay) is generally higher under fixed standard contracts, which provides us with a rational explanation for why growers prefer fixed standards contracts over tournaments. However, there is a qualification to this statement and it is that the welfare advantages of fixed standard contracts will decrease as the relative size of the common shock standard deviation increases. Thus, a ban on tournaments may enhance actual grower welfare (as opposed to ex ante welfare), but a policy maker should be aware that such bans can eliminate insurance against large common shocks. In this case, it is possible for grower welfare to be hurt by a ban.

Principal welfare was unambiguously higher with fixed performance standards when we modeled principal revenues using a simple, linear function of output. If this revenue function is at all realistic, then a ban would clearly enhance welfare for the principal. Of course, this begs the question of why a principal would ever choose a tournament in the first place if profits are higher under fixed standard contracts. A possible answer to this is that our simple revenue function may be misspecified. We therefore conducted the same analysis using a nonlinear revenue function that is based on a loss function. Our results using the nonlinear revenue function suggest that, for large common shocks, processors may earn more profits from using tournaments so that a tournament ban may negatively impact processor welfare.

In conclusion, our overall results suggest that, unless the common shock is large, a ban on tournaments may enhance both ex-post grower and processor welfare in the short run. This claim is based on our statistical findings which suggest that both agents and principals could be better off using fixed standard contracts for moderate to small common shocks. However, this leads to a key puzzle which is that, if both growers and processors prefer tournaments when common shocks are large, but prefer fixed standard contracts when common shocks are small, then why does there appear to be such political conflict in practice between processors and growers? We posit several possible explanations which can form the basis for future research. First, growers may simply not believe that tournaments are fair because they are forced to compete against each others and they are unclear about what performance standards they need to meet in any given flock or growing season. We are currently conducting experiments that investigate both subject attitudes toward experiments and what they are willing to pay to participate in fixed standard contracts versus fixed standards contracts. Second, anecdotal evidence from growers in some livestock industries suggest that processors can manipulate the inputs used by growers and/or game the tournament rankings so that there is not "an even playing field" across growers in tournament competitions. Thus, even if a well run tournament can be advantageous to growers, a tournament that includes opportunistic behavior may be economically harmful to many growers.

References

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

Cason, T.N., and Plott, C.R. "EPA's New Emissions Trading Mechanism: A Laboratory Evaluation." *Journal of Environmental Economics and Management*. 30(1996): 133-160.

Dessein, W. "Authority and Communications in Organizations," *Review of Economic Studies*. 69(2002):811-838.

Eckel, C., and N. Lutz. "Introduction: What Role Can Experiments Play in Research on Regulation?" *Journal of Regulatory Economics*. 23(2003):103-107.

Friedman, D., and Sunder, S. *Experimental Methods: A Primer for Economists*. Cambridge University Press, 1994.

Holmstrom, B. "Moral Hazard and Observability", *Bell Journal of Economics.* 10(1979): 74-91.

Holmstrom, B. "Moral Hazard in Teams" *The Bell Journal of Economics.* 13(1982):324-340

Hvide, H.K. "Tournament Rewards and Risk Taking." *Journal of Labor Economics*. 20(2002): 877-898.

Lazear, E.P. and Rosen, S. "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy,* 89(1981), pp. 841-64.

Ledyard, J. "Public Goods: A Survey of Experimental Research." In *Handbook of Experimental Economics* edited by J. Kagel and A. Roth, Princeton University Press, 1994.

Levine, H. "Robust tests for equality of variances" *Contributions to Probability and Statistics.* ed. I. Olkin, 278-292, 1960, Stanford University Press.

Levy, A. and T. Vukina. "The League Composition Effect in Tournaments with Heterogeneous Players: An Empirical Analysis of Broiler Contracts," *Journal of Labor Economics* 22(April 2004), forthcoming.

Roe, B. and S. Wu. "The Welfare Effects of Banning Tournaments When Commitment Is Impossible: Some Results from Broilers," working paper, Department of Agricultural, Environmental and Development Economics, Ohio State University, May, 2003.

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

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

Appendix A

Sample Instructions for the Tournament

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. Your pair member has an identical table.

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 both the frequency (number of pennies for each specific number) and the probability of drawing a particular 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 $0.81; if your point total is smaller than your pair member, you receive the low payment of $0.40. Whether you receive $0.81 or $0.40 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 experimenter will flip a coin to determine who 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 10th 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".

Are there any questions?

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 $0.81 as your payment
    b. Lower than your 'pair member' → Circle $0.40 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)

Appendix B

Sample Fixed Standard Contract Instructions

This experiment is identical to Experiment A in all aspects except the following.

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 $0.85 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 $0.43. Whether you receive $0.85 or $0.43 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 $0.85 as your payment
    b. Less than 41 → Circle $0.43 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)

Appendix C

Sample Auction Instructions

In the second half of today's session you will have the opportunity to earn more money by participating in two more experiments identical to the two experiments played in the first half of today's session only without the initial, non-paying practice rounds. That is, the rules and the number of paying rounds for the experiments played in the second half will be exactly like those played in the first half.

For each experiment, however, only 10 of you will be allowed to participate. Which 10 of you will participate in each experiment will be decided as follows.

For Experiment A you will fill out a Experiment A bid card. On this card you will place your chair number and the maximum number of dollars you would be willing to pay from your experimental earnings today in order to participate. You will then fill out a similar card for Experiment B. The total amount of your bids for Experiment A and Experiment B combined cannot exceed the running total of dollars you have earned so far in the experiment.

We will collect the Experiment A and Experiment B bid cards from all participants and rank them from highest to lowest for each experiment. The top 10 bids for each experiment will be allowed to play in that additional experiment.

Each participant that gains entry into an additional experiment will have his/her running dollar total decreased by the amount of the 10^{th} place bid for that experiment. Note: if your bid is higher than the 10^{th} place bid, you will pay less than the amount you bid. In other words, you will gain no advantage by bidding less than your true value for entry to the additional experiment, since it is unlikely you would have to pay the full amount you bid.

The two people with the lowest bids for each experiment will not be allowed to play in the additional experimental session and will not have any dollars deducted from their running total. They must sit quietly as the additional experiment is played.

Review of Instructions

19. Write maximum amount you are willing to pay to play an additional round of Experiment A on the Experiment A Bid Card
20. Write maximum amount you are willing to pay to play an additional round of Experiment B on the Experiment B Bid Card
21. Verify the sum of bids for Experiment A and B are not greater than your net running total.
22. The top 10 bidders for each experiment will play in an additional experiment.
23. Your net running totals will be reduced by the amount of the 10^{th} place bid if you were one of the top 10 bidders for that experiment.
24. Your net running total will not be reduced if you are not in the top 10 bidders, but you can't play in the additional experiment.