Optimal Contest Design When Common Shocks are Skewed: Theory and Evidence from Lab and Field Experiments*

John A. List
University of Chicago and NBER

Daan van Soest
VU University Amsterdam and Tilburg University

Jan Stoop
Tilburg University

Haiwen Zhou
Old Dominion University

12 March 2010

Abstract

Contests are ubiquitous. The critical link between labor hiring and promotion practices, elections, research and development, grading of students, government procurement contracts, many sporting events, and the like is that the economic structure relates rewards to *relative* performance. This fact has lead to an explosion of social science research over the past several decades clarifying the problem of incentives when competitors are rewarded according to relative performance. Yet, many first order questions remain open. One such exploration concerns the relationship between optimal contest design and the common uncertainty component. This paper begins by showing that the assumed shape of the common uncertainty component is critical in this regard: if the form of uncertainty that characterizes the tournament process is skewed, then equilibrium effort levels depend crucially on the number of competitors. As a first test of our theory we utilize a lab experiment, where important features of the theory can be exogenously imposed. We proceed to execute a field experiment, where we rely on biological models complemented by economic models to inform us of the relevant theoretical predictions. In both cases we find that the theory has a fair amount of explanatory power. More generally, from a methodological perspective our study showcases the benefits of combining data from both lab and field experiments to test economic theory.

JEL: C91, J33, C72
Key words: theory of tournaments, experiments, incentives, uncertainty

*Thanks to seminar participants at Harvard University, Princeton University, and the University of California-Santa Barbara for helpful comments. Jonathan Alevy, Stef van Kessel, Roel Ikink, and Michael Price provided excellent research assistance.

1. Introduction

Contests have a long and storied past. From the first Olympiad in 776 B.C., to Medieval England, where tournaments were an essential part of military life, to modern labor markets, contests have served an important social and economic function. While some form of contest arrangement has by now permeated nearly every walk of life, scholars have only recently begun to explore rigorously the theoretical and empirical underpinnings of various tournament formats. Over the past several decades, for instance, important theoretical work has clarified the problem of incentives when competitors are rewarded according to relative performance (for early work, see e.g., Lazear and Rosen (1981), Holmstrom (1982), Carmichael (1983), Green and Stokey (1983), Nalebluff and Stiglitz (1983), Malcomson (1984) and O'Keefe et al. (1984)).[1]

Empirically testing the theoretical predictions of contest models has taken two quite distinct paths: regression-based methods that focus on outputs and laboratory experiments that are able to measure inputs. One clever illustration of the former method is due to Ehrenberg and Bognanno (1990), who find strong support consistent with the notion that golfers on the PGA tour respond to the level and structure of prizes in tournaments. The seminal laboratory experiment is due to Bull et al. (1987), who find that effort levels converged to theoretical predictions in aggregate, but that individual effort level choices were quite noisily distributed around the equilibrium prediction.

To date, the literature has identified the prize structure and monitoring system (stochastic element) as the key means to provide the correct incentives in contests. In this

[1] The literature has provided several reasons for employing contests, including reducing monitoring costs, dealing with indivisible rewards, and minimizing risks from common uncertainties, but contests do entail potential dangers. For example, they may elicit an incorrect level of effort (moral hazard) or induce the wrong agents to participate (adverse selection).

study, we enhance the principal's choice set by showing that the number of contestants allowed to compete also has important effects on individual effort levels. It seems that the importance of the size of the contest has been left unexplored in the literature due to the standard assumption that outcomes depend on effort and a uniform stochastic component. Yet, in many settings the common uncertainty component might be skewed. For example, any zero sum game with positive or negative externalities might yield a non-uniform uncertainty component. In R&D contests, for instance, positive spillover effects might induce an increasing density function characterizing the uncertainty component. Furthermore, in many settings negative shocks might be correlated across agents, such as in labor markets.

Our theory highlights that if the form of uncertainty that affects outcomes is skewed, then the number of competitors allowed in the competition has a critical influence on equilibrium effort levels. In particular, as the size of the tournament expands, a contestant's equilibrium effort level i) decreases if the form of uncertainty has a decreasing density, ii) remains the same with a uniform density, and iii) increases if the form of uncertainty has an increasing density. The intuition is that the marginal benefit from committing effort critically relies on both the shape of the density and the number of competitors.[2]

Our first test of the theory utilizes a laboratory experiment. By studying experimental markets that differ only in the shape of the common uncertainty component, we are permitted a unique insight into whether our theoretical predictions are played out

[2] There are only a few studies of tournaments that we are aware of that vary group size. Orrison et al. (2004) study tournaments among three or more contestants, but they assume that the uncertainty component is uniformly distributed. Harbring and Irlenbusch (2003) do not use an uncertainty component at all.

in a controlled environment. Experimental methods in the lab thus permit us to study effects that would be quite difficult to identify in naturally-occurring data. In addition, it provides a useful benchmark for making inference from our field data. Similar to the results in Bull et al. (1987), we find that aggregate effort levels converge to theoretical predictions, but there is substantial noise at the individual level. In terms of comparative statics, we find mixed support for our theory when we assume that contestants are risk neutral. When we relax the risk neutrality assumption and allow risk aversion, however, our comparative static predictions are met with much greater frequency: contestants' effort levels respond predictably to changes in the number of competitors.

Our second empirical investigation—a field experiment—continues to rely on randomization as an instrument to test our theory, but proceeds in a slightly different spirit. Whereas in our laboratory experiment we impose all of the underlying assumptions of the theory and explore effort choices, under our field approach we ask whether these results continue to obtain in a field setting, where simplifying assumptions are not guaranteed to hold. Importantly, however, to test our theory it is necessary to have an observable measure of individual effort and a firm grasp of the underlying common uncertainty component.

Our search for an appropriate environment concluded when we obtained an agreement with a Dutch commercially-run recreational fishing outfit.[3] Agents in this environment commonly compete in tournaments, and we can measure their effort levels in a straightforward, non-intrusive, manner. As such, our simple experimental manipulations are viewed as natural by participants, and with the spatial arrangement of

[3] We are grateful to Ad and Thea van Oirschot of "De Biestse Oevers", Biest-Houtakker, The Netherlands, for allowing us to use their ponds for experimentation.

competitors around the lake we have priors on the shape of the common uncertainty component. This is so because at the beginning of each period we place a fixed number of Rainbow Trout into the pond. Such an introduction of fish results in a negatively skewed common uncertainty component for two reasons: the number of fish in the pond is finite and Rainbow Trout is a species that biology has taught us typically school (Liao et al. 2003). Overall, we find evidence consonant with our theoretical predictions—as the number of competitors increases, individual effort levels decline.

We view our results as having import in several domains. For instance, an important element of contest design that has not been explored rigorously in a theoretical and empirical sense is the optimal size of the tournament. Intuition in some circles is that larger tournaments induce greater levels of competition and therefore greater effort levels. Alternatively, some argue that the probability of winning a small tournament is larger, providing workers with greater incentives to exert effort (see Harbring and Irlenbusch (2003), and Orrison et al. (2004) for useful discussions). Our theory provides intuition into why such insights might not be contradictory, and provides a direction into interesting positive and normative implications heretofore not discussed. In doing so, we not only add a tool to enhance mechanism design, but provide insights into current policy debates. For example, if effort is not necessarily decreasing in the number of contestants, then merger and acquisitions cannot be justified by the argument that concentration is necessary to give firms incentives to conduct research and development.

Methodologically, both our lab and field experiments permit a glimpse of individual effort levels, which is rarely achieved in the literature on tournaments using naturally-occurring data, which have necessarily focused on outputs, rather than inputs (see, e.g, O'Reilly et al. (1988), Ehrenberg and Bognanno (1990), Main et al. (1993), Orszag (1994), Becker and Huselid (1992), and Lynch and Zax (1998, 2000), Eriksson (1999)). More generally, our study showcases the benefits of combining lab and field experimental data to test economic theory.

The remainder of our paper is organized as follows. Section 2 provides our theoretical model and experimental design. Section 3 discusses our empirical results. Section 4 concludes.

2. Tournament Theory

The theoretical literature on tournaments represents a rich lot with several interesting implications. Much of the literature in the area of labor economics can be traced to the work of Lazear and Rosen (1981), who originally clarified the problem of incentives when competitors are paid on a relative basis. Their theory lends structure to several real-world phenomena, including salary structure in corporations and payouts in sporting events. Green and Stokey (1983) pushed the argument in an important direction by demonstrating that the optimally-designed tournament dominates other reward systems when a sufficiently diffuse common shock exists.

Malcomson (1984) later highlighted certain properties of tournaments by examining incentives in an asymmetric information environment. Proceeding in a somewhat different dimension, Okeefe et al. (1984) take the structure as given and model the problem as one of intensive and extensive optimality: eliciting the correct level of effort and inducing the correct people to participate. In a labor setting, they argue that with the proper use of monitoring probability and prize structure the moral hazard and adverse selection problems can be solved.[4]

A. Theoretical model: risk neutrality

Following Zhou (2002), we assume there are n risk neutral contestants exerting effort to produce output, $n \geq 2$. In the typical triangular prize format, the agent who produces the highest level of output wins the contest and receives a reward of W_1. Each of the remaining agents receives a payoff of $W_2 < W_1$. Let μ_i denote a representative agent i's effort level, and μ_k denote the effort of her *k*th rival, $i, k \in \{1, 2, ... n\}$. Let ε_i and ε_k denote identically and independently distributed random variables which have a distribution function denoted by *F*. The distribution function is assumed to be continuous and twice differentiable and the corresponding density function is *f*. The realized output q_i of contestant i is defined as

$$q_i = \mu_i + \varepsilon_i . \tag{1}$$

Under these conditions, for q_i to be the highest level it is necessary that

$$\mu_i - \mu_k + \varepsilon_i > \varepsilon_k \text{ for all } k \neq i .$$

Assuming symmetry, all rivals' effort is the same, denoted by μ. Given the effort level of her rivals, contestant i's probability of producing the best output is $F^{n-1}(\mu_i - \mu + \varepsilon_i)$ for a given ε_i. Integrating over all possible realizations of ε_i, contestant i's expected probability of winning the contest is

[4] In a related literature spawned by Schumpeter (1950), who argued that some concentration in an industry is necessary to provide firms with sufficient incentives to invest in R&D, an abundance of studies have examined the role of contest design on equilibrium effort levels. In an early and influential work, Loury (1979) studied the relationship between market structure and R&D spending. Assuming random variables

$\int_{-\infty}^{+\infty} F^{n-1}(\mu_i - \mu + \varepsilon_i) f(\varepsilon_i) d\varepsilon_i$. Let $C(\mu_i)$ denote contestant i's cost of effort level μ_i: we assume $C' > 0$ and $C'' > 0$. Thus her expected payoff is

$$W_1 \int_{-\infty}^{+\infty} F^{n-1}(\mu_i - \mu + \varepsilon_i) f(\varepsilon_i) d\varepsilon_i + W_2 \{1 - \int_{-\infty}^{+\infty} F^{n-1}(\mu_i - \mu + \varepsilon_i) f(\varepsilon_i) d\varepsilon_i\} - C(\mu_i).$$

Contestant i chooses μ_i to maximize the expected payoff. Assuming an interior solution, the first order condition for contestant i's profit maximization is

$$(W_1 - W_2) \int_{-\infty}^{+\infty} (n-1) f(\mu_i - \mu + \varepsilon_i) F^{n-2}(\mu_i - \mu + \varepsilon_i) f(\varepsilon_i) d\varepsilon_i - C'(\mu_i) = 0.$$

In a symmetric equilibrium, $\mu_i = \mu$ and the above equation reduces to

$$(W_1 - W_2) \int_{-\infty}^{+\infty} (n-1) f^2(\varepsilon_i) F^{n-2}(\varepsilon_i) d\varepsilon_i - C'(\mu) = 0. \tag{2}$$

Using integration by parts, we find that

$$\int_{-\infty}^{+\infty} (n-1) f^2 F^{n-2} d\varepsilon = f(+\infty) - \int_{-\infty}^{+\infty} F^{n-1} f' d\varepsilon. \tag{3}$$

In (3), $f(+\infty)$ does not depend on n. As a result, we have

$$d \int_{-\infty}^{+\infty} (n-1) f^2 F^{n-2} d\varepsilon / dn = \int_{-\infty}^{+\infty} (-\ln F) F^{n-1} f' d\varepsilon. \tag{4}$$

From (2) and (4), when ΔW is fixed, the sign of $d\mu / dn$ is the same as the sign of f'. Thus, a first proposition follows:

Proposition 1: *The form of uncertainty characterizing the tournament affects the relationship between the number of contestants and equilibrium effort levels.*

The intuition behind this result is as follows. When an agent chooses her effort level, she naturally compares the marginal benefits and costs of effort. When the number of contestants increases, the probability of one or more other contestants receiving a very

are exponentially distributed, Loury (1979) showed that R&D spending decreases with the number of competing firms.

good draw is increasing, and this holds independent of whether the density function is increasing, decreasing, or uniform. The increase in this probability has two competing effects influencing the marginal benefit function. The first effect is that "pure luck" (a good realization of a contestant's random variable) is less likely to determine the winner. The larger the number of contestants, the more likely it is that at least some agents end up with high realizations (for a given distribution), and hence the more important effort is in determining the winner. The second effect is that each individual contestant's probability of having the best luck decreases. With convex effort costs, for a given prize the net marginal benefit of effort falls.

Three natural examples are intuitively plausible. First, suppose the density function is increasing. The contestant knows that she has a high probability of receiving a good draw, but she also knows that the probability of one or more other contestants receiving a good draw is increasing in group size. Hence, the larger the group, the closer the contestants are in terms of likely outcomes—good draws. As a result, the first effect dominates the second when the number of contestants increases, and effort plays an important role in selecting the winner.

Second, suppose the density function is decreasing. The contestant knows that her probability of receiving a good draw is small, whereas the probability of at least one other contestant receiving a good draw increases in group size. Hence, the larger the number of contestants, the smaller the likelihood that putting in extra effort will pay off, and hence the second effect dominates the first. Finally, as is typical in the literature, when the density function is assumed to have a zero slope, the first effect exactly cancels the second effect and effort does not change with the number of contestants.[5] Hence, in equilibrium, the following comparative statics naturally follow:

Hypothesis 1*: When contestants are risk neutral*, a *contestant's effort decreases* (if $f' < 0$)*, remains the same* (if $f' = 0$)*, or increases* (if $f' > 0$) *as the number of contestants increases.*

These predictions are empirically testable, and will form the basis of our empirical tests below.

B. Theoretical model: risk aversion

One potentially important assumption in the theory, and indeed the bulk of contest theory in general, is that agents are risk neutral. In this sense, any empirical test represents a joint hypothesis test—risk neutrality and equilibrium play. In an effort to extend this aspect of the literature, we consider our theoretical predictions when agents are risk averse. This appears to be a natural assumption, as recent explorations of individual risk preferences (e.g., Holt and Laury, 2002) suggest that a majority of agents act in a manner consistent with a model of risk-aversion when confronted with choices of lottery payoffs that are typical in lab experiments.

[5] A simple numerical example facilitates interpreting the results. Consider the case where there are only two possible outcomes; one can have either a good draw (with probability p) or a bad draw (with probability $1-p$). Furthermore, assume that the costs of effort and the size of the prize are such that it is not profitable for a contestant to exert any effort if she herself receives a bad draw and at least one other contestant receives a good draw. That means that effort only plays a role in selecting the winner (i) if either the contestant receives a good draw *and* at least one other contestant also receives a good draw (which happens with probability $p(1 - (1-p)^{n-1})$), or (ii) if the contestant receives a bad draw *and* none of the other contestants receives a good draw (the probability of which equals $(1-p)^n$). So the probability that effort matters is $\pi(n,p) = (1-p)^n + p(1 - (1-p)^{n-1})$. If n increases, $(1-p)^{n-1}$ goes to zero. If one gets a bad draw (which happens with probability $(1-p)$), the chance of winning goes to zero if group size increases. If one gets a good draw, the probability of no other contestant receiving the good draw goes to zero if group size increases. Hence, for larger groups, effort is less (more) likely to determine the winner if p is relatively small (large) – that is, if the probability mass on the good outcome is small (large). Indeed, noting that $\pi(\infty,p) = p$, we have $(\pi(2,p),\pi(\infty,p))$ equals (0.68,0.20), (0.50,0.50) and (0.68,0.80) for p=0.20, p=0.50, and p=0.80, respectively.

Continuing with our theoretical model described above, we denote the well-behaved utility function for an agent as U: $U'>0$, $U''<0$. An agent's expected utility is therefore:

$$U(W_1 - C(\mu_i))\int_{-\infty}^{+\infty} F^{n-1}[(\mu_i - \mu) + \varepsilon_i] f(\varepsilon_i) d\varepsilon_i$$
$$+U(W_2 - C(\mu_i))\{1 - \int_{-\infty}^{+\infty} F^{n-1}[(\mu_i - \mu) + \varepsilon_i] f(\varepsilon_i) d\varepsilon_i\}.$$

In a symmetric equilibrium, the condition for profit maximization becomes

$$V \equiv ((U(W_1 - C) - U(W_2 - C))\int_{-\infty}^{+\infty} (n-1) f^2(\varepsilon_i) F^{n-2}(\varepsilon_i) d\varepsilon_i$$
$$-\frac{1}{n}[U'(W_1 - C) + (n-1)U'(W_2 - C)]C'(\mu_i) = 0. \tag{6}$$

From (2) and (6), whether a contestant will commit more or less effort when they are risk averse depends on the following inequality

$$W_1 - W_2 > \frac{n(U(W_1 - C) - U(W_2 - C))}{U'(W_1 - C) + (n-1)U'(W_2 - C)}. \tag{7}$$

Since the utility function is concave, (7) always holds. Thus, ceteris paribus, risk averse agents will commit less effort than risk neutral agents. The intuition behind this result is as follows. Given that the utility function is concave, $U(W_1) - U(W_2)$ is smaller than $W_1 - W_2$; thus, as the reward to winning decreases, a contestant's effort decreases.

From (6), the relationship between a contestant's effort and the number of contestants can be explored

$$\frac{d\mu}{dn} = -\frac{\partial V / \partial n}{\partial V / \partial \mu}. \tag{8}$$

From the second order condition for a contestant's payoff maximization, $\partial V / \partial \mu$ is always negative. Thus, $d\mu / dn$ has the same sign as $\partial V / \partial n$.

Partial differentiation of (6) yields

$$\frac{\partial V}{\partial n} = [U(W_1 - C) - U(W_2 - C)]\int_{-\infty}^{+\infty} f' F^{n-1}(-\ln F)d\varepsilon$$

$$-\frac{1}{n^2}[U'(W_1 - C) + U'(W_2 - C)]C'. \qquad (9)$$

We are therefore able to derive two further hypotheses for empirical testing:

Hypothesis 2*: A risk averse contestant will commit less effort than a risk neutral contestant.*

Hypothesis 3*: If $f' < 0$ or $f' = 0$, then a risk-averse contestant's effort decreases as the number of contestants increases.*

For the increasing density case, $f' > 0$, a risk-averse agent's equilibrium effort level is ambiguous over changes in the number of rivals. As noted above, for $f' > 0$, an agent's effort level increases with the number of contestants when they are risk neutral. With the impact of risk aversion, the relationship is ambiguous because risk aversion serves to decrease the utility payoff from winning. In much the same manner, for $f' = 0$, a risk neutral contestant's equilibrium effort level does not change as the number of contestants changes, but introducing risk aversion causes this relationship to become negative.

3. Experimental Evidence

To examine the hypotheses posed above, we proceed in two distinct, but complementary, directions. We begin by imposing the major assumptions of our theory in a laboratory experiment, allowing us a crisp look into the effects of alternative common certainty components on individual effort levels. We proceed to an environment—a field experiment—where we can only be certain that a few of the critical assumptions are met: those that provide enough structure to provide theoretical predictions. The proceeding discussion will be organized as an explanation of each experimental method followed by the empirical evidence drawn from that approach.

A. Lab Experiment

We use a combination of the literature and our theory to guide the choice of experimental parameters for our lab experiment. For example, to ensure comparability with the extant literature, wherever possible we follow Bull et al. (1987) in crafting our experimental protocol. Accordingly, we use the quadratic cost function given by $C(\mu)=\frac{3}{10000}\mu^2$, where μ denotes the chosen effort level. In addition, our theory pinpoints that the reward to the winner should be \$4.50, and the reward to other contestants is \$2.40. The rewards are chosen so that subjects are provided significant incentives to exert effort and the second order condition for a contestant's payoff maximization is satisfied regardless of whether there are two or eight contestants. Also, and in the spirit of the literature, in each treatment subjects are confronted with a sequence of twenty rounds in which they are to choose an effort level in each round over the interval [0,100].

For our lab experiment we recruited subjects from the undergraduate student body at the University of Maryland. We use a between-subject design wherein each subject plays 20 rounds within one out of the six tournament treatments. Table 1 presents these six treatments, summarized as a 3X2 full factorial experimental design. Table 1 also provides sample sizes for each treatment, and presents the Nash equilibrium effort levels assuming that all participants in a group are either risk neutral or have the same level of risk aversion. Rows represent whether the treatment was carried out with 2 or 8 competitors in the contest, and columns denote the shape of the common uncertainty component: decreasing ($f' < 0$), uniform ($f' = 0$), or increasing ($f' > 0$). In the table, and in the discussion below, we denote the two-person decreasing density tournament treatment as D2; other treatments use similar acronyms.

As illustrated in the experimental instructions contained in Appendix I, subjects were informed of the number of competitors, of the structure of payoffs and costs, as well as of the fact that the tournament is to be played directly after all subjects have made their choices for that round. Thus, subjects played rounds one by one. Subjects were thus aware of the fact that they and their playing competitors could earn either $4.50 or $2.40 in each round. Finally, subjects were informed that monies earned will be summed and paid in private at the end of the experiment.

Upon completion of the tournament experiment, instructions and a decision sheet were handed out for a second experiment. This second experiment was designed to estimate subjects' risk preferences. In this part of the session, the low-payoff treatment of Holt and Laury (2002) was used—see Appendix II for the experimental instructions.[6] The treatment is based on ten choices between paired lotteries. The paired choices are included in Appendix II. The payoff possibilities for Option A, $2.00 or $1.60, are much less variable than those for Option B, $3.85 or $0.10, which was considered the risky option. The odds of winning the higher payoff for each of the options increased with each decision, and the paired choices are designed to determine degrees of risk aversion.

[6] We elected to use the low-payoff treatment of the Holt and Laury experiment to measure risk preference as the domain of earnings for this treatment [$0.10 to $3.85] approximates the equilibrium domain of per period earnings for our contest markets. We also collected data for a higher-payoff treatment of the Holt and Laury experiment where the domain of earnings more closely approximates the equilibrium domain of earnings at the session level in our contest markets. In what follows, we report only the empirical results for risk preference based upon individual response to the low-payoff Holt and Laury design. However, all tests and qualitative results are robust to the use of response to the higher-payoff experiment.

Holt and Laury (p. 1649) provide a table that will be used to categorize subjects' CRRA and CARA risk preference levels based on their ten decision choices.

After the instructions were read and all questions were answered, subjects were asked to complete their decision sheets by choosing either A or B for each of the ten decisions. Subjects were instructed that one of the decisions would be randomly selected *ex post* and used to determine their payoffs. Part of a deck of cards was used to determine payoffs, cards 2-10 and the Ace to represent "1". After each subject completed his or her decision sheet, a monitor would approach the desk and randomly draw a card twice, once to select which of the ten decisions to use, and a second time to determine what the payoff was for the option chosen, A or B, for the particular decision selected. After the first card was selected, it was placed back in the pile, the deck was reshuffled, and the second card was drawn. For example, if the first draw was an Ace, then the first decision choice would be used, and the subject's decision, A or B, would be circled. Suppose the subject selected A in the first row. The second draw would then be made. If the Ace was drawn, the subject would win $2.00. If a card numbered 2-10 was drawn, the subject would win $1.60. The subjects were therefore aware that each decision had an equal chance of being selected.

After all the subjects' payoffs were determined, they combined their payoffs from both experiments to compute their final earnings. The final payoffs were then verified against records maintained by a monitor, and subjects were paid privately in cash for their earnings. Each of the sessions took approximately 75 minutes.

Prior to moving to a discussion of the experimental results, a few noteworthy aspects of our experimental design merit further consideration. First, data were gathered in 17 sessions. Second, no student competed in more than one treatment. Third, no participation fee was used, but subjects earned on average roughly $50. Fourth, each of the treatments was carried out via linked computers; we used the Z-tree program (Fischbacher 1999).

Experimental Results of the Lab Experiments

Table 2 presents a summary of the experimental results. In this summary, we have averaged individual play over all 20 periods, effectively providing one observation per person in an effort to be as conservative as possible in our beginning statistical testing. Yet, our results are robust to using observations from rounds 11-20 only, or from rounds 15-20 only. In panel A of Table 2, the raw data show that, in aggregate, effort in the 2-player contest is consistently higher than effort levels in the 8-player contest, regardless of the shape of the random density component. This is not consonant with our risk-neutral theoretical model, which predicts that the shape of the density function critically moderates the relationship between effort and tournament size.

The observed across-the-board decrease in effort when group size is increased is consistent with our subjects being risk averse. The theory presented in Section 2 suggests that in case of negatively skewed or uniformly distributed shocks subjects in larger groups should exert less effort, whereas the impact of group size is ambiguous in case of a positively skewed shock distribution. Panels B and C of Table 2 address these predictions by categorizing the level of CRRA (and CARA) risk preference for the agent based upon estimates provided on p. 1649 of Holt and Laury (2002): risk-neutral agents are those who select 4 or fewer "safe" choices, risk averse agents are those who select more than 4 "safe" choices. This parsing permits a test of both hypotheses 2 and 3.

The first conclusion we can derive from this parsing is that the observed deviations from our theory in the aggregate data might be due to the fact that more than half (100 out of 162) of our participants reveal themselves to be risk averse rather than risk neutral. Upon splitting these types into a separate category, we find that the predictive power of our theory improves substantially. For instance, consonant with our theory, effort levels of risk-neutral and risk-averse subjects differ considerably. We therefore state a first result:

Result 1. *Risk averse players commit less effort than risk neutral contestants in all treatments.*

Evidence for this result comes from a series of Mann-Whitney rank-sum tests of treatment differences. The rank-sum test is a standard nonparametric test that has a null hypothesis of no treatment effect, or that the two samples are derived from identical populations. To construct the test statistics, we first calculated the individual mean effort levels across the twenty rounds and then ranked subjects via these means. The test statistic is normally distributed, and in our case this approach provides several p-values. We summarize the empirical results in panel D of Table 2. Comparing within tournament size and particular shapes of the stochastic component, we find that five of the six p-values are below 0.02, and for the U8 treatment we can reject the null of no difference at the $p = 0.07$ level. Taken together, this represents strong support for *Hypothesis 2.*

Whereas *Result 1* relates to the correlation of risk preferences and effort levels, we can also explore the impact of group size for given risk preferences, providing insights into *Hypothesis 3*. Comparing the observed effort levels for risk-neutral subjects for groups of 2 to effort levels of players in groups of 8 in panel B of Table 2, we find that their effort levels are largely insensitive to group size. Risk-averse subjects, however, consistently choose lower effort levels when confronted with more contestants; see panel C of Table 2. This yields our next result:

Result 2: *Whereas effort levels of risk neutral subjects do not change with changes in group size, our theory is supported in the case of risk averse contestants as their effort levels decrease for* $f' < 0$ *and* $f' = 0$ *as the number of contestants increases.*

We again turn to a series of Mann-Whitney rank-sum tests of treatment differences to provide empirical support for this result—see the relevant p-values regarding the impact of group size in panel E of Table 2. Summary results for risk neutral players show that effort levels actually increase for $f' < 0$, but decrease for $f' > 0$ and $f' = 0$. For $f' > 0$, the change is marginal, but the change is nearly significant for $f' = 0$ at conventional levels. Overall, these results suggest that effort levels among risk neutral players are largely unresponsive to changes in group size, at odds with *Hypothesis 1*. Alternatively, our theory has a fair amount of explanatory power for risk averse agents. For example, for both $f' < 0$ and $f' = 0$, we find that individual effort levels decrease significantly with increases in the number of competitors. For the $f' < 0$ case at the $p < .01$ level, for $f' = 0$ at the $p < .06$ level. Furthermore, even though theory is silent on the impact of group size in case of $f' > 0$, we find that effort levels decrease if the distribution of shocks is positively skewed for this case as well.

Another aspect of our theory contains point estimates. Although the experimental approach might not be the best tool to provide immutable point estimates (see, e.g., Levitt and List, 2007), it is instructive to consider how close our model comes to predicting actual play; cf. Tables 1 and 2. Interestingly, consonant with Bull et al. (1987), the point estimates for most treatment cells are close to the level predicted by our theory in aggregate, but not at the individual level. This is largely due to accurate aggregate predictions, but there is high variability across individuals. This variability leads to most of our theoretical predictions being within one standard deviation of the observed point estimates obtained. This is remarkable especially because the theory is based on the assumption that risk preferences are homogenous within groups. To give an example, the predicted level for risk averse agents for a given density and a group size is derived by solving the model assuming that each risk averse agent is matched with either 1 or 7 risk averse agents. In practice, groups are heterogeneous in terms of risk preferences, as group formation is random.

Although the analysis of the raw data provides evidence to support certain aspects of our theory, there has been little effort to examine the (temporal dimension of) individual data. To rectify this shortcoming, we estimate a model whereby we regress the individual effort choice on a series of treatment dummy variables, while allowing for unobservable subject- and time-effects:

$$E_{\mathrm{it}} = \mathrm{X'}_{\mathrm{it}}\beta + \varepsilon_{\mathrm{it}}, \quad i = 1,\ldots, N, \;\; t = 1,\ldots,T. \tag{10}$$

E_{it} is the effort choice for subject i in period t; X_{it} includes a full set of treatment dummies and period dummies, all of which are exogenous. In our regression estimations, the standard errors are clustered at the subject level. This allows the errors to be correlated over the periods for a subject and permits different variances and covariances across subjects. Since our theory provides no guidance on the structure of learning as captured by our time effects, we allow for maximum flexibility by using time dummies; as such, estimated trial effects can be linear, quadratic, cubic, or any other shape the data dictate.

The regression results of this OLS model are presented in Table 3.[7] Column (i) of Table 3 contains the regression results using all observations, whereas columns (ii) and (iii) present those of the risk-neutral and risk-averse subjects respectively. Regressions are run without an intercept, so that a full set of treatment dummies can be included. To show the impact of group size for each of the three cases, we included dummy variables capturing the three cases (negatively, uniformly, and positively skewed shock distributions) as well as these dummies interacted with group size dummies (for n=8).

Column (i) of Table 3 indicates that effort levels are smaller if groups are larger (significant at $p < 0.05$) for decreasing and uniform distributions, but not for increasing density functions. Even though theory predicts for the latter that effort should be significantly higher in larger groups, it is reassuring to find that when controlling for the temporal pattern in a regression analysis effort no longer falls (cf. the average effort levels presented in panel A of Table 2). Columns (ii) and (iii) yield essentially the same conclusions as presented in *Result 2*.

B. Field Experiment

As a next test of the model's predictions, we take our theory to the field where many of the theoretical assumptions cannot be assured to be met. Recently, a rich assortment of tournament studies in field settings have arisen, shedding important insights on relevant economic models. These studies revolve around estimating empirical models using naturally-occurring data, and exclusively deal with variables concerning outputs rather than inputs. For example, in sport settings, Ehrenberg and Bognanno (1990) and Orszag (1994) report golf scores as a proxy for effort. Becker and Huselid (1992) report speed and outcomes in auto racing, while Lynch and Zax (1998) report

[7] Our results are unaffected when running Tobit models with censoring at effort levels 0 and 100.

outcomes of a horse race to measure effort in a tournament. Evidence from tournaments within firms uses a similar approach: O'Reilly et al. (1988) use measures of sales, profits and number of employees to explain CEO wages; Main et al. (1993) extends this line of work in a similar spirit. Eriksson (1999) takes a different approach, but one that similarly measures outputs, as he assumes effort is equal to average profits divided by sales.

Rather than focusing on naturally-occurring data, we move this literature in a new direction by making use of a field experiment, wherein we can measure inputs. In doing so, it is important to craft an experimental design that exogenously varies our major treatment variable—number of competitors—in an environment that permits an understanding of the other important features of the situation. This approach provides us with an opportunity to observe behavior of agents who have endogenously selected into the market, while simultaneously making use of controls afforded by an experiment. To this end, we strived to exploit a naturally-occurring environment whereby the random stochastic component takes a shape that is well understood by the participants.

Finding such an environment is not trivial, but our search concluded when the operator of a recreational fishing outfit in The Netherlands agreed to provide i) access to fishermen and ii) space on the ponds to carry out our experiment. The outfit owns three rectangular fishing ponds, each of which is roughly 8500 square feet. The normal procedure is that customers pay an entry fee of 12.50–15 Euros and fish for a period of 4 to 5 hours, depending on the season. The entry fees for these ponds vary as a function of the type and number of stocked fish – rainbow trout or salmon trout. Further, customers do not have 'property rights' regarding the fish that are thrown in on their behalf; rather, at each pond there is no constraint on catch.

This setting has several features that are ideal for our purposes. First, it provides us with a participant pool that naturally competes in tournaments, and indeed some of our subjects have participated in national Dutch fishing competitions. Second, the number of fish in the pond is limited, and decreases within each tournament round as more and more fish are caught, which implies a decreasing density function of luck. Third, biological models have taught us that trout fish school, suggesting that good luck for one fisherman means bad luck for fishermen far removed on the pond—thus reinforcing the idea that our field experiment captures the case of a decreasing density function.

Yet, given that such models are developed under normal conditions—the trout being observed in their natural environment—it is important to consider whether this data pattern is observed for fish that are newly introduce to a foreign setting. As a robustness test, we explored whether the spatial correlation of caught fish is consonant with a schooling model. Our empirical results, which are available upon request, are strongly consistent with the schooling hypothesis.

Fourth, the fishing technology is geared towards exploiting the behavior of prey by continuously casting and reeling. Bait is thus dragged through the water, seducing the trout to chase and take. Although theoretically reeling in too quickly means that the trout is outrun, as we show below the number of fish caught per time period is an increasing function of the number of casts in that period. This suggests that we have a measure of effort—the number of casts per period—that is a useful measure to test our theory. Clearly, the same casting frequency may imply very different effort levels for different subjects. Thus, to account for skill and fishermen heterogeneity our design must be careful to provide within–subject treatment variability.

With these advantages, of course, come disadvantages. For example, whereas standard tournament theory is static, our field setting is dynamic in the same spirit as the empirical studies using naturally-occurring data cited above. In addition, there is natural heterogeneity in the population, whereas our model imposes symmetry. Finally, the information conditions necessary for the theory are not necessarily met in our field experiment—such as common knowledge of costs. Although these differences are not exhaustive, they highlight that field experiments present a tradeoff: they give up some of the controls of a laboratory experiment (such as induced valuations, or robots guaranteed to play equilibrium strategies against human subjects) in exchange for increased realism. In this manner, our field experiment matches the real-world settings which tournament theory attempts to explain: our fishermen are not told explicitly the distributions of other's valuations and they have previous experience in this environment. In this manner, the exploration provides a useful middle ground between the tight controls of the laboratory and the vagaries of completely uncontrolled field data.

The execution of the tournament experiment was straightforward and followed four steps. First, sports fishermen were recruited via a registration list during the week previous to the planned session. Second, upon arrival, we explained the experimental instructions in a quiet area removed from the other customers. As shown in the experimental instructions in Appendix III, we explained that we rented a specific pond and that each subject will participate in 4 tournaments with an alternating number of competitors—either 1 or 7 per tournament. Every tournament lasts exactly one hour, and the winner of a tournament is the person who catches the most fish during the hour. The anglers were told that the winning prize is 10 euro's, independent of group size. In case of a draw, the winner is determined by whoever caught the first fish. In case of a tie, we flip a coin to determine the winner.[8]

Third, shortly before the first tournament we stocked the pond with 58 rainbow trout, and at the beginning of each round we replenished the stock that was taken out. During the stocking process, we allocated initial fishing spots, which were determined via a random process—drawing a numbered spot tag from a closed bag. Before each new tournament began, all participants drew a new numbered spot tag and were re-allocated on the pond. Fishing only occurs on the long sides of the pond, and we always use 8 of the 10 fishing spots on each side. A whistle blow marks the beginning and end of each tournament.

The fourth and final step involves participant remuneration. As noted above, the agent received 10 Euro's for each tournament victory, and hence the maximum prize money per subject is 40 Euro's. In addition, each subject received 5 Euro's participation fee. As a final bit of compensation, we collected all fish and redistributed them lump-sum to session participants.[9] The pecuniary outlay for the experiment, consisting of the costs of fish and the payments to the fishermen, was roughly 2,100 Euro's.

Before discussing our empirical results, a few outstanding issues merit brief elaboration. First, we ran 3 sessions with 16 participants each session; subjects were

[8] We were careful to follow the rules applied by the Dutch Trout Fishing Championships except for the tie breaker. In the official Championships the total weight of the fish caught determines the winner in case of a tie. This is not feasible in our experiment since each participant is in four tournaments. Breaking ties on the basis of total weight of fish caught in the particular tournament round would imply using four coolers per fishermen to store the fish separately. Each sports fisherman usually just carries one, and hence for practical purposes we used the time elapsed before catching the first fish as a tie breaker.

[9] One might have chosen to not allow participants to take home any fish. We wished to avoid waste (in total 487 fish were caught), and could not give them to a charity due to perishability. We therefore decided to redistribute them equally among all participants in a session. The marginal incentive to catch another fish (apart from the increased likelihood of winning the tournament) is thus 1/16th of its value. Since this marginal incentive is small and independent of treatment we opted for this approach.

allowed to compete in only one session each. Second, given that a within-subject design was necessary, we alternated the group sizes of the tournaments within each session. In sessions 1 and 3 the tournaments were played in group sizes of 2, 8, 2 and 8 in rounds 1-4 respectively, and in session 2 group sizes were 8, 2, 8 and 2. Thus, in session 1 we had eight tournaments of n=2 in rounds 1 and 3 and two tournaments of n=8 in rounds 2 and 4. Session 3 yielded identical samples sizes. For session 2, we have two tournaments of n=8 in rounds 1 and 3, and eight tournaments of n=2 in rounds 2 and 4. Each subject was always aware of the number of competitors in her tournament, as well as of the identity of his/her competitor(s). Finally, in light of our theoretical model and the stochastic component, under our design we have one comparative static prediction to test: *contestant's effort decreases as the number of contestants increases*. This prediction should hold for both risk neutral and risk averse competitors, hence we do not gather data on individual risk posture.

Experimental Results

Before formally testing our theoretical prediction, we need to show that there is a positive relationship between effort intensity and the number of fish caught. Our approach is to build a panel data regression model around an ordered probit regression model:

$$Y_{it}^{*} = X_{it}'\beta + \varepsilon_{it}, \tag{11}$$

where Y_{it}^{*} is the number of fish caught by individual *i* in period *t*, X_{it} is a vector of person-specific exogenous variables, with the main variable being effort intensity. Effort intensity is the number of casts per minute corrected for the time elapsed between fish caught and the moment at which the fisherman restarts fishing.[10] Also included in X_{it} is location on the pond, and the random error component ε_{it} either includes subject fixed effects, or we cluster standard errors at the subject level. Finally, β is the estimated response coefficient vector. Although we do not directly observe the true Y_{it}^*, we do observe an approximation of Y_{it}^*:

$$Y_{it} = 0 \text{ if } Y_{it}^* \leq 0; \ = 1 \text{ if } 0 < Y_{it}^* \leq \phi_1; \ = 2 \text{ if } \phi_1 < Y_{it}^* \leq \phi_2; \ \ldots; \ = 9 \text{ if } \phi_8 < Y_{it}^* \leq \phi_9. \quad (12)$$

The ϕ_i are unknown parameters that are estimated jointly with β. As such, we obtain threshold levels of the marginal value of effort by measuring how exogenous variable vector X_{it} affects fish caught.[11]

Empirical results are presented in Table 4. In both specifications, it is clear that the number of fish caught is increasing in effort intensity. For example, the empirical model in column (i), which accounts for the data dependencies by clustering standard errors, suggests that the positive relationship holds at the $p < .05$ level. Likewise, the empirical model in column (ii), which accounts for the data dependencies by including subject specific fixed effects, also suggests the positive relationship holds at conventional levels. As a robustness test, in Appendix IV we include estimates from two count data models, which both show that the number of fish caught is increasing in effort intensity.

---

[10] That means that the denominator is 60 minutes minus the time needed to land all fish caught, to get them off the hook, and to put new bait on the hook.

[11] A few aspects of our estimation procedure merit further consideration. First, since the ϕ_i's are free parameters, there is no significance to the unit distance between the set of observed values of Y, thus avoiding symmetric treatment of one-unit changes in the dependent variable. Second, estimates of the marginal effects in the ordered probability model are quite involved because there is no meaningful conditional mean function. We therefore compute the effects of changes in the covariates on the j probabilities: $\partial \text{Prob}[\text{cell } j]/\partial X_i = [f(\phi_{j-1} - X_i'\beta) - f(\phi_j - X_i'\beta)]*\beta$; where $f(\bullet)$ is the standard normal density, and other variables are defined above. By definition, these effects must sum to zero since the probabilities sum to one. These estimates are available upon request.

Having provided support for the conjecture that effort and the number of fish caught are positively correlated, we now turn to our main interest of measuring the impact of group size on fishing effort. Upon doing so, we are able to state our final result:

Result 3: *Consonant with our theory, there is some evidence that subjects decrease their effort intensity when the number of competitors increases.*

As a first glimpse into the received data patterns, we provide Table 5. Panels A and B in this table include means and standard deviations of the effort levels across treatment categorized by period, without adjusting for the data dependencies – for sessions 1+3 and session 2, respectively. Panels C and D of Table 5 provide the number of agents who change their effort levels in response to an increase or decrease in the number of competitors in accordance with our theory. Two data patterns are generally observed. First, there is an important starting point effect: effort levels tend to decline after period 1, and significantly so. Second, accounting for this trend, there is a tendency for effort levels to be higher in the 2-person tournaments, consonant with our theory, albeit more clearly so in session 2 than in sessions 1+3. Panel B shows that in session 2 effort increases significantly if group size is decreased for the second time (i.e., between rounds 3 and 4). Panel C shows that 56.25% of the participants in sessions 1 and 3 increase their effort when group size is decreased for the first time (i.e., from round 2 to round 3), but only 50% decreases effort when group size is decreased again (from round 3 to round 4). The results are again stronger for session 2 – see panel D. When group size is increased for the first time (from round 2 to round 3), 68.75% of the participants decrease their effort level, whereas no less than 81.25% of the subjects increase their effort levels when group size is decreased for the second time (between rounds 3 and 4).

As a more formal test, we follow equation (10) and estimate a panel data model at the individual level. Given the shortness of the panel, we use first differences and analyze changes in effort as follows:

$$dE_{it} / E_{it} = \beta_1 * Decreasegroupsize_{it} + \phi_t + \varepsilon_{it} \qquad (13)$$

where E_{it} is the effort choice for subject i in period t, dE_{it} is the difference in effort of subject i between period t and period $t - 1$, Decreasegroupsize$_{it}$ is a dummy variable with value 1 if group size decreases from 8 in period t-1 to 2 in period t, and 0 otherwise. Furthermore, ϕ_t is a set of three time dummies, indicating a 1 when moving from one period to another. Finally, ε_{it} are standard errors clustered at the subject level. In this model, the dummy Decreasegroupsize$_{it}$ captures decreases in periods 2-3 in Sessions 1 and 3 and in periods 1-2 and 3-4 in Session 2, leaving the period dummies free to capture responses in case group size increases (in periods 1-2 and 3-4 in Sessions 1 and 3, and in period 2-3 in Session 2).

The results of this analysis are presented in column (i) of Table 6. We find evidence consonant with our theory: effort intensity increases when group size decreases – but the result is not strongly significant. We also find that effort intensity decreases by 23% between rounds 1 and 2 (as evidenced by the coefficient on Period 1-2), but is essentially constant thereafter. This is consistent with our earlier assertion of novelty wearing off, resulting in a substantial drop in effort when moving from tournament 1 to tournament 2 but in a much smaller drop in effort (or even no drop in effort) when group size is changed for the second and third time.

We explore the consequences of removing the novelty effect by omitting the first change (from period 1 to 2; see column (ii) in Table 6) or even the first two changes (from period 1 to 2, but also from period 2 to 3; see column (iii) in the table) from our data set. In column (ii) we find that decreases in group size increase effort intensity by just over 10% (p-value = 0.058), while the sign of the period transition dummies indicate that effort falls if group size increases. And we find a comparable percentage change in column (iii) (with a p-value of 0.074). We conclude that the field experiment provides moderate support for our major theoretical prediction.

4. Conclusions

Designing optimal incentive schemes is perhaps one of man-kinds oldest activities. From the Dead Sea Scrolls to scribes on tombs of ancient kings, rudimentary and clever incentive structures to motivate a particular course of action have been extolled.[12] For their part, economists have produced a rich assortment of models that lend insights into the various factors that are likely to influence equilibrium market behavior. In a work environment, for example, the models teach us that individual effort levels are critically tied to the incentive scheme in place.

Such explorations have naturally led scholars to clarify the problem of incentives when competitors are rewarded according to relative performance. Although much progress as been made in the past several decades, the relationship between optimal contest design and the common uncertainty component remains under explored. In this paper, we expand the literature in this direction by providing a new theory as well as experimental evidence testing our theory.

[12] One of Aesop's fables provides an example in point: "A father, being on the point of death, wished to be sure that his sons would give the same attention to his farm as he himself had given it. He called them to his bedside and said, "My sons, there is a great treasure hid in one of my vineyards." The sons, after his death, took their spades and mattocks and carefully dug over every portion of their land. They found no treasure, but the vines repaid their labor by an extraordinary and superabundant crop."

The theory provides several predictions, including how risk aversion influences equilibrium play, and that the assumed shape of the common uncertainty component is critical in determining equilibrium effort levels. In this regard, if the form of uncertainty that characterizes the tournament process is skewed, then equilibrium effort levels depend crucially on the number of competitors. In the case of a uniformly distributed common uncertainty component, the number of competitors is not predicted to influence effort levels.

Our line of attack to test the theory is a two-pronged variant, though the empirical approaches are complementary in nature. Our first method is to use a laboratory experiment, which permits us to study markets that differ only in the shape of the common uncertainty component, allowing a unique insight into whether changes in the component's shape itself can lead to predicted changes in behavior. Lab experimental methods thus allow us to study such effects that would be difficult to identify in naturally occurring data. Our second approach is to maintain randomization, but design an experiment in the field that resembles the important features of our theory and permits us to examine effort levels directly.

Overall, the lab data provide mixed support for our theory when we assume that contestants are risk neutral. When we relax the risk neutrality assumption and allow risk aversion, however, our comparative static predictions are met with much greater frequency. Furthermore, our model accurately predicts how risk aversion influences equilibrium play. The field data complement these insights by providing evidence consonant with the theory within a special case of the theory—when the common uncertainty component is negatively skewed. In this case, we find some evidence that adding competitors decreases individual effort levels.

We view our results as having import in several circles. For instance, it provides a theoretical basis for the disparate views concerning the optimal number of players in a contest, and clarifies when larger tournaments should induce greater levels of effort. Such insights might aid the contest designer interested in optimal wage schemes, government procurement contracts for R&D contests, company promotional policies, and optimal mechanism design more generally.

Methodologically, this study showcases that by controlling the type of uncertainties characterizing the contest process, a crisp view of the impact of the number of contestants on a contestant's effort can be achieved. Likewise, by controlling for the number of competitors, one can estimate the effects of changing the nature of the uncertainty component. Combining these insights with data patterns from a field experiment permits one to make much stronger inference than one could with either in isolation. This is so because our field experiment can check the robustness of laboratory results in a natural setting, where the mathematical assumptions of the theory cannot necessarily be guaranteed to hold. This approach provides a useful middle ground between the controlled environment of the laboratory and the unruly nature of uncontrolled field data.

References

Becker, Brian E., and Huselid, Mark A. 1992. “The incentive effects of tournament compensation systems.” *Administrative Science Quarterly* 37, pp. 336-350.

Bull, Clive, Schotter, Andrew and Weigelt, Keith. 1987. “Tournament and piece rates: an experimental study.” *Journal of Political Economy* 95, pp.1-31.

Carmichael, Lorne. 1983. “Firm-specific human capital and promotion ladders.” *Bell Journal of Economics* 14, pp.251-258.

Ehrenberg, Ronald, and Bognanno, Michael. 1990. “Do tournaments have incentive effects?” *Journal of Political Economy* 98, pp. 1307-1324.

Eriksson, Tor. 1999. “Executive compensation and tournament theory: Empirical tests on Danish data.” *Journal of Labor Economics* 17, pp. 262-280.

Fischbacher, Urs. 1999. z-Tree: Zurich Toolbox for Readymade Economic Experiments: Experimenter's manual. Institute for Empirical Research in Economics, University of Zurich.

Green, Jerry, and Stokey, Nancy. 1983. “A comparison of tournaments and contracts.” *Journal of Political Economy* 91, pp. 349-364.

Harbring, Christine and Irlenbusch, Bernd. 2003. “An experimental study on tournament design.” *Labour Economics* 10, pp. 443-464.

Holmstrom, Bengt. 1982. “Moral hazard in teams.” *Bell Journal of Economics* 13, pp. 324-340.

Holt, Charles and Laury, Susan. 2002. “Risk aversion and incentive effects.” *American Economic Review* 92(5), pp. 1644-1655.

Lazear, Edward and Rosen, Sherwin. 1981. “Rank-order tournaments as optimum labor contracts.” *Journal of Political Economy* 89(5), pp. 841-864.

Levitt, Steven D. and List, John A. 2007. “What do laboratory experiments measuring social preferences reveal about the real world?” *Journal of Economic Perspectives* 21(2), Spring, pp. 153-174.

Liao, James C., Beal, David N., Lauder, George V., and Triantafyllou, Michael S. 2003. “Fish exploiting vortices use less muscle.” *Science* 302, pp. 1566-1569.

Loury, Glenn. 1979. “Market structure and innovation.” *Quarterly Journal of Economics* 93(3) August, pp. 395-410.

Lynch, James G., and Zax, Jeffrey .S. 1998. “Prizes, selection and performance in Arabian horse racing.” *Discussion Paper*, University of Colorado.

Lynch, James G., and Zax, Jeffrey. S. 2000. “The rewards to running: prize structure and performance in professional road racing.” *Journal of Sports Economics* 1; pp. 323-340.

Main, Brian G.M., O’Reilly, Charles A., and Wade, James. 1993. “Top executive pay tournaments of teamwork.” *Journal of Labor Economics* 11, pp. 606-628.

Malcomson, James. 1984. “Work incentives, hierarchy, and internal labor market.” *Journal of Political Economy* 92, pp. 486-507.

Nalebuff, Barry and Stiglitz, Joseph. 1983. “Prizes and incentives: toward a general theory of compensation and competition.” *Bell Journal of Economics* 14(1), Spring, pp. 21-43.

O’Keeffe, Mary, Viscusi, W. Kip and Zeckhauser, Richard. 1984, “Economic contests: comparative revenue schemes.” *Journal of Labor Economics* 2, pp. 27-56.

O'Reilly III, Charles A. and Main, Brian G. and Crystal, Graef S. 1988. “CEO compensation as tournament and social comparison: a tale of two theories.” *Administrative Science Quarterly* 33(2) Jun., pp. 257-274.

Orrison, Alannah, Schotter, Andrew and Weigelt, Keith. 2004. “Multiperson tournaments: an experimental examination.” *Management Science* 50(2) Feb., pp. 268-279.

Orszag, Jonathan M. 1994. “A new look at incentive effects and golf tournaments.” *Economics Letters* 46, pp. 77– 88.

Schumpeter, Joseph. 1950. *Capitalism, socialism, and democracy*, Third edition, New York: Harper &Row.

Zhou, Haiwen. 2002. *Three essays on industrial organization*, Ph. D. Dissertation, University of Maryland, College Park.

Table 1 – Design of the lab experiment

	Density		
	Decreasing f'<0	Uniform f'=0	Increasing f'>0
2-player contest	D2; *n* = 30	U2; *n* = 32	I2; *n* = 28
Prediction Risk Neutral	53.08	43.75	53.08
Prediction Risk Averse	28.92	23.84	28.92
8-player contest	D8; *n* = 24	U8; *n* = 24	I8; *n* = 24
Prediction Risk Neutral	26.94	43.75	73.47
Prediction Risk Averse	14.68	23.84	40.03

Notes: Entries represent treatment acronym, sample sizes and theoretical predictions assuming all subjects in a tournament are either risk-neutral, or risk-averse. Treatment U8 denotes a uniform density 8-player contest. Twenty-four subjects participated in this treatment. Risk neutral (risk averse) subjects are predicted to choose effort level 43.75 (23.84) when groups are homogenous in terms of risk preferences. Each subject in every treatment participated in 20 rounds.

Table 2 – Lab results

	Density		
	Decreasing $f' < 0$ Avg. Effort (std. dev.)	Uniform $f' = 0$ Avg. Effort (std. dev.)	Increasing $f' > 0$ Avg. Effort (std. dev.)
A. All players			
2-player contest	42.13(12.49)	37.88(17.13)	41.54(16.76)
	$n = 30$	$n = 32$	$n = 28$
8-player contest	33.38(18.01)	27.05(15.76)	35.22(21.02)
	$n = 24$	$n = 24$	$n = 24$
Comparison (p-values, MW)	0.020	0.011	0.0517
B. Risk Neutral players (RN)			
2-player contest	48.31(8.67)	48.30(16.99)	53.74(9.96)
	$n = 15$	$n = 11$	$n = 12$
8-player contest	50.20(15.53)	37.21(10.99)	50.71(24.05)
	$n = 9$	$n = 5$	$n = 10$
C. Risk Averse players (RA)			
2-player contest	35.96(12.91)	32.42(14.81)	32.38(14.98)
	$n = 15$	$n = 21$	$n = 16$
8-player contest	23.29(10.21)	24.38(15.95)	24.15(8.07)
	$n = 15$	$n = 19$	$n = 14$
Comparisons (p-values, MW)			
D. Risk attitudes			
RN 2-player vs RA 2-player	0.012	0.019	0.000
RN 8-player vs RA 8-player	0.001	0.070	0.004
E. Group size			
RN 2-player vs RN 8-player	0.612	0.089	0.644
RA 2-player vs RA 8-player	0.003	0.0598	0.0417

Notes: Standard deviations are in parenthesis after average effort, and n denotes the number of observations. Comparisons are done by a Mann-Whitney test for independent samples. For example, the difference between risk neutral players and risk averse players in a 2-player contest is statistically significantly different with a p-value of 0.012.

Table 3 – OLS estimation results of effort levels in the lab

Variable	All subjects	Risk neutral subjects	Risk averse subjects
Decreasing	47.858***	49.165***	44.701***
	(2.822)	(3.973)	(3.773)
Decreasing*(n=8)	-8.752**	1.893	-12.670***
	(4.270)	(5.434)	(4.151)
Uniform	43.601***	49.159***	41.157***
	(3.400)	(5.285)	(3.922)
Uniform*(n=8)	-10.824**	-11.090	-8.037*
	(4.366)	(6.691)	(4.812)
Increasing	47.260***	54.600***	41.122***
	(3.567)	(4.159)	(4.398)
Increasing*(n=8)	-6.319	-3.032	-8.231**
	(5.263)	(7.860)	(4.226)
Time fixed effects	Yes	Yes	Yes
N	3240	1240	2000
R^2	0.628	0.765	0.555

Notes: Dependent variable is subject's effort choice. Decreasing is a dummy variable which has a value of 1 if the distribution of the random shock is decreasing. The dummy variables Uniform and Increasing have a similar meaning. Furthermore, (n=8) is a dummy variable which has a value of 1 if the group size of the subject is 8. Standard errors are robust and clustered at the subject level and are reported in parenthesis below the coefficient estimates. ***, **, Significant at the 1%, and 5% respectively.

Table 4 – Ordered Probit estimation results for catch of fish in the field

Variable	(i) Ordered probit, clustered se	(ii) Ordered probit
Effort intensity	0.928**	3.618***
	(0.445)	(0.416)
Quadrant fixed effects	Yes	Yes
Subject fixed effects	No	Yes
N	192	192
LogPseudoLikelihood	-366.81	-314.22
Pseudo-R^2	0.023	0.163

Notes: Dependent variable is subject's catch of fish in a period. Effort intensity is the number of casts per minute, corrected for the time elapsed between fish caught and the moment at which a fisherman restarts fishing. Robust standard errors are reported in parenthesis below the coefficient estimates. In column (i), the standard errors are clustered at the subject level. Column (ii) has subject fixed effects. ***, **, Significant at the 1%, and 5% respectively.

Table 5 – Field results

	Period				Change in group size (GS)		
	1	2	3	4	Wilcoxon Signed-Rank Test		
A. Session 1+3	2-player	8-player	2-player	8-player	GS increases (first time)	GS decreases (first time)	GS increases (second time)
Effort	0.723	0.646	0.641	0.672	p=0.020	p=0.963	p=0.507
St.dev.	(0.239)	(0.248)	(0.210)	(0.277)			

	Period						
	1	2	3	4	Wilcoxon Signed-Rank Test		
B. Session 2	8-player	2-player	8-player	2-player	GS decreases (first time)	GS increases (first time)	GS decreases (second time)
Effort	0.704	0.563	0.537	0.610	p=0.001	p=0.352	p=0.023
St.dev.	(0.225)	(0.170)	(0.202)	(0.212)			

	Percentage of subjects that changes effort in accordance with the theory		
C. Session 1+3	GS increases (first time)	GS decreases (first time)	GS increases (second time)
	68.75%	56.25%	50%
D. Session 2	GS decreases (first time)	GS increases (first time)	GS decreases (second time)
	20%	68.75%	81.25%

Notes: Effort intensity is the average number of casts per minute, corrected for the time elapsed between fish caught and the moment at which a fisherman restarts fishing. Wilcoxon Signed-Rank Test is for within person change of effort, where n = 32 for session 1 + 3 and n = 16 for session 2.

Table 6 – OLS estimation results of effort levels in the field

Variable	(i)	(ii)	(iii)
Decreasegroupsize	0.043	0.105*	0.121*
	(0.056)	(0.055)	(0.066)
Period 1-2	-0.231***		
	(0.058)		
Period 2-3	-0.08	-0.120**	
	(0.051)	(0.056)	
Period 3-4	0.001	-0.014	-0.020
	(0.048)	(0.045)	(0.051)
Number of observations	144	96	48
R-squared	0.169	0.056	0.056

Notes: Dependent variable is a subject's percentage change in effort between periods. Decreasegroupsize is a dummy variable which has a value of 1 if group size decreases from 8 in period *t*-1 to 2 in period *t*. Period 1-2 is a dummy variable which has a value of 1 when a subject moves from period 1 to period 2. The dummy variables Period 2-3 and Period 3-4 have similar meanings. Standard errors, robust and clustered at the subject level, are reported in parenthesis under the coefficient estimates. ***, **, * Significant at the 1%, 5%, and 10% respectively.

Appendix I. Experimental Instructions for D2 Treatment

Introduction

This is an experiment in decision making. The instructions are simple; if you follow them carefully and make good decisions, you could earn a considerable amount of money, which will be paid to you in cash.

First Part of the Experiment

As you read these instructions you will be in a room with a number of other subjects. Each subject has been randomly assigned an ID number.

The first part of the experiment consists of 20 decision rounds. In each decision round you will be grouped with another subject by a random drawing of ID numbers. That subject will be called your "group member." Your group member will remain the same throughout the entire experiment. The identity of your group member will not be revealed to you and your identity will not be revealed to him or her.

Experimental Procedure

In the experiment you will perform a simple task. Attached to these instructions is a sheet called your "Decision Costs Table." This sheet shows 101 numbers from 0 to 100 in the first column. These are your Decision Numbers. All subjects have the same "Decision Costs Table".

Associated with each Decision Number on the Decision Costs Table are Decision Costs. Note that the higher the Decision Number chosen, the greater is the associated cost.

Your computer screen should look like the following:

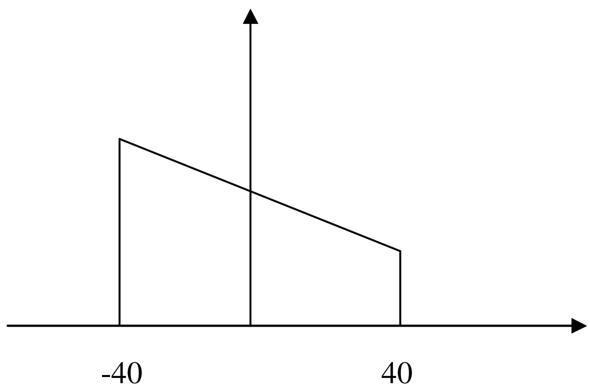
PLAYER#: DECISION#: RANDOM#: TOTAL#: EARNINGS

In each decision round the computer will ask each subject to choose a Decision Number. Therefore, you and your group member will each separately choose one Decision Number. Using the number keys at the top of the keyboard, you will enter your selected number and then hit the Return (Enter) key. To verify your selection, the computer will then ask you the following question:

Is your number ___? [Y/N]

If you want to select the number shown as your Decision Number, hit the Y key. If not, hit the N key and the computer will ask you to select a number again. You do not need to hit the Return key after entering Y or N. After you have selected and verified the Decision Number, this number will be recorded on the screen in Column 2, and its associated cost will be recorded in Column 5.

After you have selected your Decision Number, the computer will ask you to generate a random number. You do this by hitting the space bar (the long key at the bottom of the keyboard). Hitting the bar causes the computer to select one of the 81 numbers that fall between -40 and +40 (including 0). The figure below illustrates the random number generator. The figure reveals that THE HIGHER THE NUMBER, THE LOWER THE PROBABILITY THAT THIS NUMBER WILL BE CHOSEN WHEN YOU HIT THE SPACE BAR. For example, the probability that the computer selects, say, +40, is smaller than the probability that the computer selects +17. For another example, the probability that the computer selects –10 is smaller than the probability that the computer selects –37.



Each subject faces the same task and random number generator. Hence, each subject will follow the same procedure, so that each subject generates his or her random number separately. After you hit the space bar, the computer will record your random number on the screen in Column 3.

Calculation of Payoffs

Your payment in each decision round will be computed as follows. After you select a Decision Number and generate a random number, the computer will add these two numbers and record the sum on the screen in Column 4. We will call the number in column 4 your "Total Number." After computation of the Total Number, the computer will compare your Total Number with your group member's Total Number. On the basis of this comparison, the computer will tell you whether you receive the "Fixed Payment" 4.5 or the "Fixed Payment" 2.4.

Recall that there are 2 members in your group. If your Total Number is higher than your group member, then you will receive the Fixed Payment 4.5. Otherwise, you will receive the Fixed Payment 2.4. If both you and your group member have the same Total Number, the computer will randomly decide whether you or your group member gets the higher Fixed Payment. Think of this procedure as though the computer is assigning “heads” to one group member, “tails” to the other, and then flipping a coin. If “heads” turns up, the group member assigned “heads” receives the high Fixed Payment.

Whether you receive the high Fixed Payment or the low Fixed Payment depends only on whether your Total Number is greater than the Total Number of your group member. IT DOES NOT DEPEND ON HOW MUCH GREATER IT IS.

The computer will record (on screen in Column 6) which Fixed Payment you receive. If you receive the high Fixed Payment (4.5), then “M” will appear on Column 6. If you receive the low Fixed Payment (2.4), then “m” will appear.

After indicating which Fixed Payment you receive, the computer will subtract your associated Decision Cost (Column 5) from this Fixed Payment. This difference represents your earnings for the round. The amount of your earnings will be recorded on the screen in Column 6, right next to the letter (“M” or “m”) showing your Fixed Payment. The earning of your group member will be calculated in exactly the same way.

Continuing Rounds

After Round 1 is over, you will perform the same procedures for Round 2, and so on for 20 rounds. In each round you will choose a Decision Number (of course, you may choose the same one in different rounds), you will again generate a random number by pressing the space bar, your Total Number will be compared to your group member’s Total Number, and the computer will calculate your earnings for the round.

Note that the Decision Cost subtracted in Column 5 is a function only of the Decision Number that you selected; i.e., your random number does not affect the amount subtracted. Also, note that your earnings depend on the following: the Decision Number you select (both because it contributes to your Total Number and because it determines the amount to be subtracted from your Fixed Payment), the Decision Number your group member selects, your generated random number, and your group member’s generated random number.

When round 20 is completed, the computer will ask you to press any key on the keyboard. After you do this, the computer will add your earnings from each of the 20 rounds. After completion of these 20 rounds we will move to the second part of the experiment. After this second part, we will then pay you the amount you earn from both parts of the experiment.

Appendix II: Experimental Instructions for Risk Aversion Experiment

Record your subject number from the previous part on your decision sheet. Your decision sheet shows ten decisions listed on the left. Each decision is a paired choice between OPTION A and OPTION B. You will make ten choices and record these in the final column, but only one of them will be used in the end to determine your earnings. Before you start making your ten choices, please let me explain how these choices will affect your earnings for this part of the experiment.

We will use part of a deck of cards to determine payoffs; cards 2-10 and the Ace will represent "1". After you have made all of your choices, we will randomly select a card twice, once to select one of the ten decisions to be used, and a second time to determine what your payoff is for the option you chose, A or B, for the particular decision selected. (After the first card is selected, it will be put back in the pile, the deck will be reshuffled, and the second card will be drawn.) Even though you will make ten decisions, only one of these will end up affecting your earnings, but you will not know in advance which decision will be used. Obviously, each decision has an equal chance of being used in the end.

Now, please look at Decision 1 at the top. OPTION A pays $2.00 if the Ace is selected, and it pays $1.60 if the card selected is 2-10. OPTION B yields $3.85 if the Ace is selected, and it pays $0.10 if the card selected is 2-10. The other decisions are similar, except that as you move down the table, the chances of the higher payoff for each option increase. In fact, for Decision 10 in the bottom row, the cards will not be needed since each option pays the highest payoff for sure, so your choice here is between $2.00 or $3.85.

To summarize, you will make ten choices: for each decision row you will have to choose between OPTION A and OPTION B. You may choose A for some decision rows and B for other rows, and you may change your decisions and make them in any order. When you are finished, we will come to your desk and pick a card to determine which of the ten decisions will be used. Then we will put the card back in the deck, shuffle, and select a card again to determine your money earnings for the OPTION you chose for that decision. Earnings for this choice will be added to your previous earnings, and you will be paid all earnings in cash when we finish.

So now please look at the empty boxes on the right side of the record sheet. You will have to write a decision, A or B in each of these boxes, and then the card selection will determine which one is going to count. We will look at the decision that you made for the choice that counts, and circle it, before selecting a card again to determine your earnings for this part. Then you will write your earnings in the blank at the bottom of the page.

Are there any questions? Now you may begin making your choices. Please do not talk with anyone else while we are doing this; raise your hand if you have a question.

Table A1 – The ten paired lottery-choice decisions with low payoffs

Option A	Option B	Expected payoff difference
1/10 of $2.00, 9/10 of $1.60	1/10 of $3.85, 9/10 of $0.10	$1.17
2/10 of $2.00, 8/10 of $1.60	2/10 of $3.85, 8/10 of $0.10	$0.83
3/10 of $2.00, 7/10 of $1.60	3/10 of $3.85, 7/10 of $0.10	$0.50
4/10 of $2.00, 6/10 of $1.60	4/10 of $3.85, 6/10 of $0.10	$0.16
5/10 of $2.00, 5/10 of $1.60	5/10 of $3.85, 5/10 of $0.10	-$0.18
6/10 of $2.00, 4/10 of $1.60	6/10 of $3.85, 4/10 of $0.10	-$0.51
7/10 of $2.00, 3/10 of $1.60	7/10 of $3.85, 3/10 of $0.10	-$0.85
8/10 of $2.00, 2/10 of $1.60	8/10 of $3.85, 2/10 of $0.10	-$1.18
9/10 of $2.00, 1/10 of $1.60	9/10 of $3.85, 1/10 of $0.10	-$1.52
10/10 of $2.00, 0/10 of $1.60	10/10 of $3.85, 0/10 of $0.10	-$1.85

Note: Taken from Holt and Laury (2002).

Appendix III: Experimental Instructions for the Field Experiment (Session 1 + 3)

IIIA: Summary of rules handed out to the participants

Tournament

- You will participate in four tournaments. You will be assigned into groups which change in composition over the day. Therefore, it is likely that you participate in tournaments with changing participants.
- The duration of each tournament is 1 hour.
- The winner of a tournament is the one who catches most fish of his/her group.
- Rainbow trout and salmon trout both count as 1 fish.
- In case of a draw, the winner is the one who caught his/her first fish first.
- Whenever you catch a fish, make sure to communicate this to the organizers behind the desk. In that way, we can make sure that we do not make mistakes in counting the number of fish caught.
- For each tournament, only the winner receives a price. He/she receives €10.
- The beginning and end of a tournament is indicated by a blow on a whistle.
- Each tournament is a separate tournament with each its own winner. It does not matter who has the most fish at the end of the day.

Sequence of events

- Each participant plays 4 tournaments, tournament A through D.
- At the beginning of a new tournament you will change your fishing spot. Between the tournaments will be a break of 5 minutes.
- In tournament A, you will play in groups of 2. You draw a fishing spot number out of a bag. The one who fishes opposite of you is the other participant of the tournament. The participants at spot 20 and spot 1 play a tournament, the participants at spot 19 and spot 2 play a tournament, and so on. The winner of each pair of participants is the one who catches most fish.
- Tournament B is played in groups of 8. You draw a fishing spot number out of a bag. One group consists of the 8 spots on the canal side of the pond (these are numbers 1 through 4 and 17 through 20); the second group consists of the 8 spots on the meadow side of the pond (these are numbers 7 through 10 and 11 through 14). The winner of each of the two groups is the one of the 8 participants who catches most fish in this tournament.
- Between tournament B and C is a break of 15 minutes.
- Tournament C is played in groups of 2 (just like tournament A). You draw a fishing spot number out of a bag. The one who fishes opposite of you is the other participant of the tournament. The participants at spot 20 and spot 1 play a tournament, the participants at spot 19 and spot 2 play a tournament, and so on. The winner of each pair of participants is the one who catches most fish.
- Tournament D is played in groups of 8 (just like tournament B). You draw a fishing spot number out of a bag. One group consists of the 8 spots on the canal side of the pond (these are numbers 1 through 4 and 17 through 20); the second

group consists of the 8 spots on the meadow side of the pond (these are numbers 7 through 10 and 11 through 14). The winner of each of the two groups is the one of the 8 participants who catches most fish in this tournament.

Putting fish into the pond

- For tournament A (which starts at 9.30 a.m.) we put 3 rainbow trout into the pond for each participant; 3 x 16 = 48 rainbow trout in total. In addition, we put in 10 extra rainbow trout. In total, 58 rainbow trout are put into the pond.
- For tournaments B, C, and D (which start at approximately 10.35 a.m., 11.50 a.m. and 12.55 p.m.) we put a number of rainbow trout into the pond equal to the total catch (both rainbow trout and salmon trout) of the previous tournament. This means that at the start of each tournament there is an equal number of fish in the pond.

Payment

- Your total earnings consist of your earnings in tournament A, B, C, and D.
- You are not allowed to keep each fish you catch! The total amount of caught fish is divided equally at the end of the day.
- For your participation you will receive €5.

IIIB: Rules read out loud by the researcher

Welcome to this study by Tilburg University. Before we start, we want to point out two things. Firstly, this study is independent of the organization 'de Biestse Oevers'. We are grateful that we are allowed to conduct this study here, but this organization has nothing to do with what we are doing here. All responsibility lies with Tilburg University. Secondly, we want to make clear that this study has nothing to do with the well-being of animals, environmental causes or the like. As researchers, we accept the rules and habits of the sports fishing as it is practiced at 'de Biestse Oevers'. We cannot tell you the exact aim of this study. We do want to stress that your privacy is protected; none of the results we report can be traced on an individual level.

As you know, you don't have to pay to take part in this study. The fishing fee is paid by Tilburg University. Each fish you catch, you are allowed to take home. In addition, you can earn money.

We ask you to abide strictly by the rules which we impose.

The study

In the next four hours, we ask you to fish according to the rules as we will explain them now. All rules that normally hold at 'de Biestse Oevers' remain in place. This means that it is not permitted to throw fish you have caught back into the pond, you are only allowed to fish with one rod, you are only allowed to use a scoop net to set fish ashore, you are only allowed to use the usual types of bait, etc.

Today you will participate in four tournaments. Each tournament takes 1 hour. The winner of a tournament is the one who catches most fish of his/her group. You are allowed to catch as much fish as possible. The pond is mainly stocked with rainbow trout, but there may also be salmon trout in the pond. Each fish you catch carries equal weight in determining who wins a tournament.

In case of a draw between two or more participants, the winner is the one who caught his/her fish in the least amount of time. In case this also results in a draw, we will toss a coin to determine the winner.

Whenever you catch a fish, please communicate this to the organizers behind the table. Wait for them to answer your call (by means of a thumb raised in the air). In this way, we make sure that we do not make mistakes in counting the number of fish caught. For each tournament there is only a prize for the winner. He/she receives €10.

The beginning and end of each tournament is marked by a whistle; each tournament lasts exactly 1 hour. At the moment the second whistle sounds, you have to you're your line and hook out of the water. If at that moment a fish is attached to your hook, you can land this fish and count it to your score.

The total duration of the study is about 4.5 hours, from 9.30 a.m. until 2.00 p.m. Each tournament is separate from the other tournaments. There is no prize for having caught the most fish at the end of the day.

You will play four tournaments. Two times you will participate in a tournament with 7 other participants (in a group of 8), and two times you will participate in a tournament with 1 other participant (in a group of 2). During the study, the composition of a group changes. The spot at which you fish is determined by means of a lottery. The first tournament, tournament A (starting at 9.30 a.m.) is played in groups of 2. You draw a fishing spot number out of a bag. The one who fishes opposite of you is the other participant of the tournament. The participant at spot 20 and spot 1 play a tournament, the participant at spot 19 and spot 2 play a tournament, and so on. The winner of each pair of participants is the one who catches most fish.

Tournament B (starting at 10.35 a.m.) is played in groups of 8 participants. You draw a fishing spot number out of a bag. One group consists of the 8 spots on the canal side of the pond (these are numbers 1 through 4 and 17 through 20); the second group consists of the 8 spots on the meadow side of the pond (these are numbers 7 through 10 and 11 through 14). The winner of each of the two groups is the one of the 8 participants who catches most fish in this tournament.

Between tournament B and C there is a break of 15 minutes.

Tournament C (starting at 11.50 a.m.) is again played in groups of 2 (just like tournament A). You draw a fishing spot number out of a bag. The one who fishes opposite of you is the other participant of the tournament. The participant at spot 20 and spot 1 play a tournament, the participant at spot 19 and spot 2 play a tournament, and so on. The winner of each pair of participants is the one who catches most fish.

Tournament D is played in groups of 8 (just like tournament B). You draw a fishing spot number out of a bag. One group consists of the 8 spots on the canal side of the pond (these are numbers 1 through 4 and 17 through 20); the second group consists of the 8 spots on the meadow side of the pond (these are numbers 7 through 10 and 11 through

14). The winner of each of the two groups is the one of the 8 participants who catches most fish in this tournament.

Stocking fish

For the first tournament, tournament A (which starts at 9.30 a.m.) we put 3 rainbow trout into the pond for each participant; 3 x 16 = 48 rainbow trout in total. In addition, we put in 10 extra rainbow trout. In total, 58 rainbow trout are put into the pond.

For tournaments B, C, and D (which start at approximately 10.35 a.m., 11.50 a.m. and 12.55 p.m.) we put a number of rainbow trout into the pond equal to the total catch (both rainbow trout and salmon trout) of the previous tournament. This means that at the start of each tournament there is an equal number of fish in the pond.

Payment

You will receive 5 euro for your participation. In addition, you will receive 10 euro for each tournament which you have won.

You are not allowed to keep all fish that you have caught. All fish caught will be divided equally among all participants at the end of a tournament.

Questions

If you have any questions regarding the rules of this study, you can ask them now, but also during the study. We do not answer questions regarding how best to fish. We also do not answer questions regarding the nature of this study.

Appendix IV: Alternative Model to Explore the Effect of Effort on Catch

A second approach to measuring the effect of effort on catch is to utilize the fact that our dependent variable is a count measure of the number of fish caught. Such a regressand is typically analyzed using a Poisson regression model, or in the case of overdispersion a negative binomial model. The Poisson model assumes that the number of fish caught for individual *i* at time *t* is drawn from a Poisson distribution with parameter λ_{it}. Consequently, the probability of observing a given number of fish caught is given by:

$$\Pr(Caught_{it} = g_{it}) = \frac{\exp\{-\lambda_{it}\}\lambda_{it}^{g_{it}}}{g_{it}!}, \quad g_{\mathrm{it}} = 0,1,2,\ldots \qquad \text{(IV.1)}$$

where $\ln(\lambda_{it}) = x_{it}\beta$, with variables defined in the text, and δ is a single, unknown parameter.

Although estimation of equation (IV.1) is straightforward via maximum likelihood, it might be the case that the variance of Caught is greater than the mean of caught. In such cases, a negative binomial model is an appropriate approach. We tested for over dispersion, and rejected the null of no over dispersion. We therefore present the results using the negative binomial approach.

Table A4 – Negative binomial estimation results for catch of fish in the field

Variable	(i)	(ii)
Effort intensity	1.642***	3.674**
	(0.701)	(0.278)
Quadrant fixed effects	Yes	Yes
Subject fixed effects	No	Yes
N	192	192
Log PseudoLikelihood	-371.03	-321.97

Notes: Dependent variable is subject's catch of fish in a period. Effort intensity is marginal effect of the number of casts per minute, corrected for the time elapsed between fish caught and the moment at which a fisherman restarts fishing. Standard errors are reported in parenthesis below the coefficient estimates. In column (i), the standard errors are clustered at the subject level. Column (ii) has subject fixed effects, while the standard errors are clustered at the session level. ***, **, Significant at the 1%, and 5% respectively.