Sabotage and the Endogenous Design of Tournaments*

Preliminary Version

Please do not quote or circulate

October 2004

Christine HARBRING♣ Bernd IRLENBUSCH♠

Abstract

If organizations implement incentive systems in which rewards depend on relative rather than on absolute performance, activities are induced from at least two dimensions: (i) *productive* activities increase the own output of an agent whereas (ii) *destructive* ones (also called sabotage) reduce the output of the competitors. As sabotage activities can barely be analyzed by collecting data from the field this paper adopts an experimental approach. We set up an experiment to analyze the influence of endogenous tournament design on behavior regarding both activity dimensions. Our main findings are: (1) effort and sabotage increase with widening the wage gap but sabotage increases to a greater extent. (2) Given a fixed prize spread the principal is able to induce an increased output by providing the agents with higher wages. (3) When participants have the possibility to communicate via email messaging principals choose high fixed wages most frequently. Interestingly, this increases efficiency which is mainly due to a decrease in sabotage activities.

JEL Classification: D23, J33, J41, L23, C72

Key Words: Contests, Relative Performance Evaluation, Personnel Economics, Experiments

---

* We thank Christian Grund, Matthias Kräkel, Edward Lazear, Reinhard Selten, and Dirk Sliwka for valuable comments. All errors are our own. We are grateful to Sebastian Kube who programmed the experimental software. Financial support by the *Deutsche Forschungsgemeinschaft* through grant KR2077/2-1 and the *European Union* through the EU-TMR Research Network *ENDEAR* (FMRX-CT98-0238) is gratefully acknowledged.

♣ University of Bonn, Department of Economics, Adenauerallee 24-42, D-53113 Bonn, Germany, christine.harbring@wiwi.uni-bonn.de

♠ Corresponding author: London School of Economics, IIM, Houghton Street, London, WC2 2AE, United Kingdom, b.irlenbusch@lse.ac.uk

I. Introduction

Reward schemes in which remuneration is based on relative rather than on absolute performance are widely recognized as an important component in the toolbox of incentive system designers for modern organizations (see e.g. GIBBONS 1998, LAZEAR 1999, PRENDERGAST 1999). The advantages that are credited to such incentive mechanisms are manifold ranging from diminishing the influence of global shocks, ex ante commitment of the employer to pay a certain amount, the sufficiency of ordinal ranking of output, mitigation of the hidden action problem, etc. However, also severe drawbacks have been identified, from which the most prominent are collusion and sabotage between agents. The latter results from the fact that agents can choose between at least two classes of activity dimensions in order to increase their payment, i.e., they cannot only intensify their productive effort but also deteriorate their competitors' performance by exerting destructive activities. Sabotage is pervasive whenever relative performance pay is encountered. For a prominent example from sports recall the Tonya-Harding-Nancy-Kerrigan case where Harding's rival Kerrigan was injured in an attack hatched by Harding's ex-husband to keep Kerrigan off the Olympic ice skating team in 1994. Or remember the famous chariot race with Charleton Heston in "Ben Hur" when he is sabotaged by his competitor which is a fictitious, but very illustrative example for sabotage. Other examples can be found in presidential election campaigns where tremendous effort is exerted to damage the other candidate's reputation via negative campaigning.[1]

Whereas the latter form of sabotage is legally allowed within limits, destructive effort in organizations has detrimental effects on output and therefore is strictly forbidden. Since recently, the interest for the potential advantages and drawbacks of competitive incentive schemes within organizations have dramatically increased in the course of the controversial debate on *forced rankings*.[2] According to estimates a quarter of the Fortune 500 companies (e.g. General Electric, Cisco, Intel, Hewlett Packard etc.) link part of the individual merit of employees to a relative performance evaluation. The idea of forced ranking schemes is that the frequency of ratings must follow a certain distribution that is determined ex ante.[3] In some

---

[1] See e.g. KONRAD (2000), HARRINGTON and HESS (1996) or SKAPERDAS and GROFMAN (1995).

[2] Of course, besides competition within companies also competition between companies for tight markets can take the form of a winner-take-all contest (e.g. FRANK and COOK 1995).

[3] See e.g. MURPHY (1992), BOYLE (2001). The forced distribution could take the form of a normal distribution: if the group of employees to be evaluated is sufficiently large one might assume that there are only a few top performers, many people whose performance is on an average level while there are few low performers. Such forced ranking systems are usually implemented if primarily subjective measures are available to force managers to differentiate their evaluation and use the whole bandwidth of grades.

cases the bottom 10% to 20% low performers who are identified via relative performance evaluation are advised to leave the company. These management practices are also known as *rank and yank*. The problem that cooperation among employees is put at risk given such incentive schemes is heavily discussed by economists as well as practitioners.[4] As examples for sabotage in firms think of any form of blocking cooperation such as actively withholding viable information, transferring false information or damaging work tools which are necessary for the work done by other employees. Thus, from a firm's point of view it appears of eminent importance to be aware of how different incentive design issues – e.g. the magnitude of the spread between wages received by agents with high and low performance – do influence the amount of sabotage being exerted.[5] In essence, agents are likely to tune their sabotage activity balancing two aspects: the competition between agents and fairness considerations within the principal-agent relationship. This interplay of fairness towards the principal and the competition between agents induced by relative performance based incentive schemes is at its very heart an empirical question. Ideally, one would wish to compare sabotage levels observed under systematically different reward systems from the field. But, unfortunately, the output destroying feature makes sabotage to be an activity which is performed in secret and hardly observable. This turns the task of collecting reliable field data on sabotage into an almost unsolvable challenge. In the present study, we opt for an experimental approach and introduce a new experimental game to investigate the influence of wage compression on sabotage activities of agents in relative performance reward systems. An experiment has the decisive advantage that one is able to exactly observe the sabotage activities without abstaining from behavior shown by real actors. Moreover, one can focus on clear-cut institutional designs. Our main experimental findings are the following:

1. *Wage Differential* – Stronger incentives may not pay off. In our setting agents respond to tournament incentives by increasing effort and sabotage while sabotage is increased to a greater extent such that in total the output does not increase with intensified incentives.
2. *Wage Sum* – Paying higher wages is profitable. Contrary to standard economic theory we find that the wage level has an impact on the agents' decisions. Effort and sabotage are higher with higher wage levels and on average this results in increased output.

---

[4] For a critical discussion of forced ranking systems see for example PFEFFER and SUTTON (1999).

[5] If the *x%* low performers identified during a relative performance measurement process with a forced ranking are supposed to leave the company the prize differential for employees appears to be quite high, e.g. in consulting firms or investment banks with an up-or-out promotion system.

3. *Communication* – Communication, though cheap talk, improves the situation for all. It is well known that the introduction of an opportunity to communicate to each other is likely to increase efficiency in social dilemma games. Because communication is a natural opportunity in real-world organizations, the analysis of the influence of communication is an important issue in our framework. Although from a theoretical perspective a flat rate does not provide any incentives for performance, fixed wages are most frequently selected by the principal. Agents respond by exerting slightly higher effort and reduce sabotage vigorously. Thus, the communication possibility enables the participants to achieve more output and higher payoffs for all players.

The rest of the paper proceeds as follows. In the next section, we give a short overview of related literature. Afterwards, we introduce and analyze a simple tournament model with two activity dimensions which serves as the baseline for our experiment. Section IV deals with the experimental setting based on the model described in the previous section. In section V, we present the experimental results and section VI concludes the paper.

II. Related Literature

The essential characteristics of payment schemes based on relative performance are captured in the well-known tournament model introduced by LAZEAR and ROSEN (1981) in which several agents compete for prizes by trying to attain the highest observable output. LAZEAR and ROSEN (1981) find that a tournament among two risk neutral agents can induce efficient effort levels like piece rates. They show that in equilibrium effort positively depends on the prize spread, i.e., the difference between winner and loser prize, which constitutes one of the fundamental results of tournament theory. In their model, however, only productive effort is considered. Following this seminal work several authors presented inspiring results regarding the incentive characteristics of tournaments, see e.g. GREEN and STOKEY (1983), NALEBUFF and STIGLITZ (1983), O'KEEFFE, VISCUSI, and ZECKHAUSER (1984), and ROSEN (1986).

Given the increased relevance of tournament-like reward schemes, surprisingly few empirical studies exist which can roughly be categorized into three classes. The first type of studies focuses on executive compensations in firms investigating the theoretically derived prediction of ROSEN (1986) who offers an explanation for the extraordinarily high pay of top managers, e.g. MAIN, O'REILLY and WADE (1993), ERIKSSON (1999). In these studies firm performance usually serves as an approximation for the productive effort of managers. The second category comprises studies that make use of tournament data sets available from sports in which, however, the tournament design is determined by the specific rules or the sports institutions, e.g. LYNCH and ZAX (1998), EHRENBERG and BOGNANNO (1990a, 1990b), ORSZAG (1994), and BECKER and HUSELID (1992). Due to the difficulties to obtain adequate data from field tournaments the third class of empirical studies evolves in the literature: several experimental studies provide valuable deepening of the understanding of the incentive effects provided by tournaments: BULL, SCHOTTER, and WEIGELT (1987) compare tournament incentives and piece rates. They also increase the prize spread but simultaneously vary other parameters, e.g. the cost function. Interestingly, average effort levels support the theoretical prediction for both schemes but effort in tournaments is much more variable. In other experimental studies, e.g. WEIGELT, DUKERICH, and SCHOTTER (1989), SCHOTTER and WEIGELT (1992), NALBANTIAN and SCHOTTER (1997), ORRISON, SCHOTTER, and WEIGELT (1997), VAN DIJK, SONNEMANS, and VAN WINDEN (2001), HARBRING and IRLENBUSCH (2003a), different aspects of tournament theory are analyzed, i.e., the influence of heterogeneity among agents, tournament size, and different prize structures. All studies mentioned so far, however, concentrate only on one single activity dimension, i.e., productive effort, while abstracting from the sabotage option.

LAZEAR (1989) is the first one who provides a theoretical analysis of tournaments when agents can exert (productive) effort *and* (destructive) sabotage. His analysis reveals that the larger the spread between winner prize and loser prize the higher are agents' activities along both dimensions. This result implies that pay compression may be optimal from efficiency considerations if due to a higher prize spread the increase in effort is outperformed by the simultaneous increase in sabotage. Thus, LAZEAR provides an argument that it may be beneficial for a firm to implement equitable pay structures. There is only one empirical study of which we are aware that is concerned with effort *and* sabotage in tournaments. In their innovative work GARICANO and PALACIOS-HUERTA (2000) investigate the effects of an exogenous change on the reward structure to gain insights in the consequences of a prize structure variation. In 1995 the FIFA (*Fédération Internationale de Football Association*) decided to increase the number of points from two to three that a team obtains for a win. The new rules were implemented worldwide in all soccer leagues. The effect of this increase of the prize spread on performance is analyzed by taking the number of forwards as a proxy for the amount of productive effort and the number of defenders as a measure of sabotage. The latter can be seen as specialists brought into play to reduce the other teams' output. In addition, the number of goals and sanctions (yellow and red cards) are included in the analysis. In this study "sabotage" is allowed, at least with respect to defend the own goal. The analysis confirms that both activities have increased with the introduction of the new reward structure, i.e., effort and sabotage are higher given the higher prize spread.[6]

Because exerting sabotage in a firm is in general strictly forbidden, gathering data on those efforts seems quite demanding not to say impossible. In experiments, however, one is able to allow agents to exert both dimensions of effort under the influence of sharply separated design features. Moreover, decisions regarding productive and destructive effort are quantifiable. Thus, experiments appear to be an appropriate tool for the analysis of behavior in tournaments with sabotage.[7] The effect of varying the prize spread has already been investigated empirically with regard to the exertion of productive effort. Up to now, however, the influence of wage compression on sabotage as well as the interaction of a principal and agents in a context where sabotage is possible lacks a thorough empirical investigation.[8] In the present study we approach this gap.

In HARBRING and IRLENBUSCH (2003b) we have already experimentally investigated a situation where a principal may endogenously select tournament prizes, and agents may exert productive and destructive activities. However, the effect of wage compression is not analyzed *ceteris paribus* in the sense that the total wage sum is not kept constant. Falk and FEHR (in progress) simultaneously and independently conducted an experiment similar to our setting. They also include the opportunity of agents to sabotage their competitors. However, their experimental design follows a slightly different research agenda. Whereas in FALK and FEHR agents are not informed about the principal's payoff function to explicitly exclude fairness considerations in our setting agents know the payoff function as well as the principal's exact payoff yielded in each round. Moreover, we allow the principal to endogenously determine pay compression *and* the total sum of wages, while FALK and FEHR keep the total sum of wages constant in each treatment.

As mentioned above, we allow participants to communicate in an additional treatment. It is known from other experimental studies that the introduction of communication can enhance cooperative behavior in social dilemma games and raise efficiency (ORBELL, DAWES and VAN

[6] In a related study DRAGO and GARVEY (1998) present evidence on the influence of incentive schemes on helping effort in work groups in Australian companies based on answers to questionnaires. They find that helping effort is reduced when promotion incentives are strong. In a sense, helping effort can be seen as opposite behavior of sabotage. Thus, one could argue that the tendency to behave destructively towards colleagues is increased by high wage dispersion.

[7] We are aware that by choosing the experimental method we are forced to boil down the real-world setting to its very essentials – as it is always the case with economic modeling. For a comprehensive overview on arguments for labor market experiments see FALK and FEHR (2003).

DE KRAGT 1988, OSTROM and WALKER 1991, GÄCHTER and FEHR 1999, CHARNESS and DUFWENBERG 2002, BROSIG, OCKENFELS and WEIMANN 2003).[9] In real-world organizations people communicate to each other, and thus, our research questions regarding the interaction of a principal and agents should also be investigated with a communication device. Following earlier experimental work indicating that efficiency can be enhanced by communication sabotage should be lower in our communication condition compared to the baseline situation.

III. A simple model of tournaments with two activity dimensions

In this section we sketch a simple two-stage game with $n + 1$ players, i.e., n agents and one principal, in which the principal selects a wage contract at the first stage before the agents exert two activities: productive effort and a sabotage activity (see Figure 1).

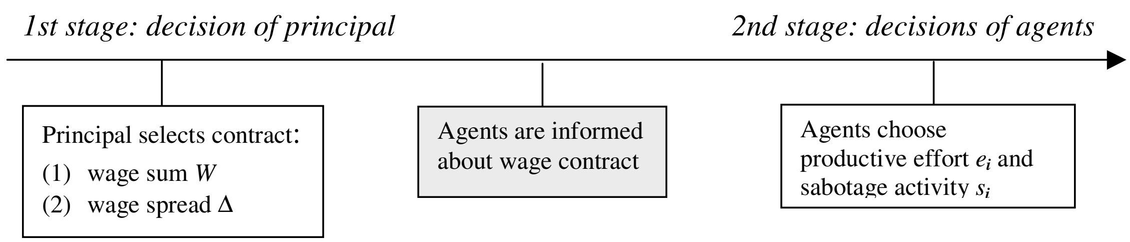

Figure 1: Sequence of the game

The principal may opt for a certain wage sum $W > 0$ and the compression of wages. In the simplest case she selects full wage compression, i.e., an equitable pay structure, and each of the n agent receives $\frac{W}{n}$. If unequal wages are to be distributed among the agents we assume that agents take part in a tournament in which they compete for a winner prize, and agents with the n-1th lowest output or less receive the loser prize. We denote the winner prize by M, the loser prize by m with ($M > m \geq 0$) and the prize differential ($M - m$) by Δ, with ($nm + \Delta$) = W as the sum of prizes has to be equal to the total sum of wages.

The strategy of an agent i is a pair (e_i, s_i) where $e_i \in [0, \ldots, \bar{e}]$ denotes an effort level and $s_i \in [0, \ldots, \bar{s}]$ is a sabotage activity which negatively influences the output of all other agents. The output y_i of agent i is determined by the following production function

[8] Recently, several investigations have shown that fairness and intentions play a quite important role in employment relationships, see e.g. FEHR, KIRCHSTEIGER, RIEDL (1993), BEWLEY (1999), CHARNESS (2000), DUFWENBERG and KIRCHSTEIGER (2000), FALK, FEHR, and FISCHBACHER (2000), FALK and GÄCHTER (2002).

[9] For an overview see also KAGEL and ROTH (1995), SALLY (1995), CRAWFORD (1998) and CAMERER (2003).

$$y_i = e_i + \varepsilon_i - \sum_{i \neq j} s_i \qquad (1)$$

with ε_i as a random variable which is uniformly distributed over the interval $[-\bar{\varepsilon}, +\bar{\varepsilon}]$ and assumed to be i.i.d. for all agents. The random component, ε_i, can be thought of as production luck or measurement error. Every agent who exerts effort or performs a sabotage activity has to bear costs, which are described by the two convex functions $C_e(e_i)$ and $C_s(s_i)$.[10] All agents have to submit their effort and sabotage decisions simultaneously. The expected payoff for agent i is given by

$$E\Pi_i(e_i, e_{-i}, s_i, s_{-i}) = F(e_i, e_{-i}, s_i, s_{-i})M + [1 - F(e_i, e_{-i}, s_i, s_{-i})]\, m - C_e(e_i) - C_s(s_i) \qquad (2)$$

with $F(e_i, e_{-i}, s_i, s_{-i})$ denoting the probability for agent i to receive a winner prize if all other agents choose effort levels $e_{-i} = (e_1, e_2, ..., e_{i-1}, e_{i+1}, ..., e_n)$ and sabotage activities $s_{-i} = (s_1, s_2, ..., s_{i-1}, s_{i+1}, ..., s_n)$.

As a benchmark let us have a look at the equilibrium behavior. For simplicity we concentrate on cost functions of the type $C_e(e_i)=e_i^2/c_e$ and $C_s(s_i)=s_i^2/c_s$ and assume that all agents are risk neutral. The expected payoff of the agents can be written as

$$E\Pi_i(e_i, e_{-i}, s_i, s_{-i}) = m + F(e_i, e_{-i}, s_i, s_{-i})\Delta - e_i^2/c_e - s_i^2/c_s \qquad (3)$$

If an interior equilibrium exists one has to consider the first order conditions

$$\frac{\partial F(e_i, e_{-i}, s_i, s_{-i})}{\partial e_i}\Delta = \frac{2e_i}{c_e} \text{ and } \frac{\partial F(e_i, e_{-i}, s_i, s_{-i})}{\partial s_i}\Delta = \frac{2s_i}{c_s}. \qquad (4)$$

Given our assumptions one can show that in a symmetric equilibrium the marginal probabilities of winning depend only on the size of the interval from which the random component in the production function is drawn (see Appendix)[11], i.e., one can show that

$$\frac{\partial F(e_i, e_{-i}, s_i, s_{-i})}{\partial e_i} = \frac{\partial F(e_i, e_{-i}, s_i, s_{-i})}{\partial s_i} = \frac{1}{2\bar{\varepsilon}}. \qquad (5)$$

---

[10] Note that we implement identical cost functions for all agents. We do not analyze the behavior of heterogeneous "personalities" like "doves" and "hawks" as in LAZEAR (1989) who models agents who differ in their marginal cost of sabotage.

[11] Proofs are given by KRÄKEL (2000). ORRISON, SCHOTTER, and WEIGELT (1997) sketch a similar argumentation as we do.

Thus, our first order conditions reduce to

$$\frac{1}{2\bar{\varepsilon}}\Delta = \frac{2e_i}{c_e} \quad \wedge \quad \frac{1}{2\bar{\varepsilon}}\Delta = \frac{2s_i}{c_s} \tag{6}$$

from which we obtain the effort level and the sabotage activity played in equilibrium

$$e^* = \frac{\Delta c_e}{4\bar{\varepsilon}} \text{ and } s^* = \frac{\Delta c_s}{4\bar{\varepsilon}}. \tag{7}$$

To ensure that an interior solution exists and that agents have no incentive to deviate to activities of zero the following condition has to be satisfied: $\Delta / n \geq C_e(e^*) + C_s(s^*)$, i.e., the expected gain of an agent must not be lower than his cost. Thus, the parameters have to be chosen such that this condition is fulfilled. Moreover, the highest possible eligible effort level must exceed the equilibrium effort level, i.e., $e^* < \bar{e}$. An analogous statement holds for the sabotage activity ($s^* < \bar{s}$).

The principal's expected payoff increases with the total effort level exerted by the agents and decreases with the total sabotage activity

$$E\Pi_P(e, s) = \tau\left[E\left(\sum_i y_i\right)\right] - \theta W = \tau\left[\sum_i e_i - \sum_i \sum_{j \neq i} s_j\right] - \theta W \tag{8}$$

where τ with $\tau > 0$ indicates the value of one unit of output for the principal and θ, with $0 < \theta \leq 1$ the proportion of wage costs the principal has to bear.[12]

Assuming that the principal may set tournament prizes we analyze how her payoff is affected by a certain pay structure. Thus, in equilibrium the principal receives the following expected payoff dependent on her selection of the two parameters, the prize spread Δ and the total wage costs *W*:

$$E\Pi_P(e^*, s^*) = \frac{\tau n \Delta}{4\bar{\varepsilon}}\left[c_e - (n-1)c_s\right] - \theta W. \tag{9}$$

---

[12] Note that the principal is modeled in a way that she bears only a proportion of the total wage costs. Thus, she can be thought of being a manager who implements the wage system and is rewarded in proportion to the output that is produced.

The principal's payoff is *ceteris paribus* increasing respectively decreasing by $\frac{\tau n}{4\bar{\varepsilon}}\left[c_e - (n-1)c_s\right]$ with widening the prize spread. Additionally, wage costs of θW are deducted from her payoff.

If $c_e > (n\text{-}1)c_s$ holds the principal's payoff increases with enlarging the prize differential for a given wage sum. This means that the output produced in equilibrium increases *ceteris paribus* with the prize spread as cost of effort is sufficiently below cost of sabotage. Thus, the additional output generated via productive effort is not overcompensated by additional sabotage activities. Finally, the principal's payoff is reduced by a fraction of the total sum of distributed wages.

In a subgame perfect equilibrium the principal anticipates the behavior of the agents and chooses a contract that maximizes her payoff. If the principal chooses full wage compression $\Delta = 0$ (*fixed wages*) rational and purely money-maximizing[13] agents should exert no activity at all. The subgame perfect equilibrium for our parameters is given in the next section.

IV. Experimental Design and Procedure

We consider tournaments with n = 3 agents and 1 principal. Table 1 depicts the design alternatives for the wage contracts. The principal chooses a wage sum $W_i \in \{W_L = 300, W_H = 600\}$[14] and selects one of the five prize differentials Δ_i with $i = 0, \ldots, 4$. A prize spread of zero is denoted by Δ_0, i.e., all players receive the same wage irrespective of the output they have achieved. The other prize spreads are positive, i.e., there is one winner prize M and two loser prizes m. The principal is also allowed not to offer any contract at all which results in a payoff of zero for the principal as well as all agents in this round.

[13] For an analysis with inequity averse players see GRUND and SLIWKA (2002) and DEMOUGIN and FLUET (2003).

[14] All payments and costs are given in "talers" which is the fictitious currency unit in the laboratory.

Table 1: Design alternatives of wage contract

	Prize differential Δ_i	Low wage level $W_L = 300$		High wage level $W_H = 600$	
No Incentive	$\Delta_0 = 0$	Fixed wage with 100 for each agent		Fixed wage with 200 for each agent	
		Loser prize *m*	Winner prize *M*	Loser prize *m*	Winner prize *M*
Incentive	$\Delta_1 = 48$	84	132	184	232
	$\Delta_2 = 96$	68	164	168	264
	$\Delta_3 = 144$	52	196	152	296
	$\Delta_4 = 192$	36	228	136	328

Table 2 depicts the theoretic effort and sabotage predictions for a given contract. After having been informed about the offered wage contract, all agents i choose their effort level e_i and their sabotage activity s_i simultaneously out of the following sets of integers: $e_i \in \{0, ..., 100\}$ and $s_i \in \{0, ..., 50\}$. The random variable ε_i as one part of the output of agent i is uniformly distributed (i.i.d. for all agents) over the integer interval [-60, + 60]. We use the parameters $c_e = 70$ and $c_s = 20$ which lead to the cost functions $C_e(e_i)=e_i^2/70$ and $C_s(s_i)=s_i^2/20$. The sabotage activity is assumed to be more expensive than productive effort as agents must exert some effort to conceal their destructive activity. After each round all players observe the output of each agent, and the agents additionally are informed about the principal's payoff.

From standard tournament theory (LAZEAR and ROSEN 1981) and the theoretic prediction derived above we know that the agents' effort and sabotage choices should depend on the prize spread but not on the wage level. Table 2 shows that the parameters are chosen such that the output is increasing in widening the prize spread. The agent's payoff is decreasing in the prize spread because of the cost of increased effort and sabotage they exert. Thus, for a given wage sum an agent's payoff is highest in a fixed-wage contract.

The value τ of one unit of output for the principal is set to $\tau = 3$ and the cost parameter to $\theta = 0.3$. Given this parameters, in equilibrium the principal's payoff increases *ceteris paribus* with the prize differential, i.e., the principal prefers the incentive contract with the highest wage differential Δ_4. Furthermore, in a subgame perfect equilibrium the principal chooses the low wage level W_L =300 because she has to bear a proportion of the total wage costs.

Table 2: Theoretic prediction for each wage contract

	No Incentive	Incentive			
	Δ_0	Δ_1	Δ_2	Δ_3	Δ_4
Effort	0	14	28	42	56
Sabotage	0	4	8	12	16
Output	0	6	12	18	24

Two treatments are conducted which are based on the same experimental setting as described above. Whereas in the *baseline* treatment agents are not allowed to communicate to each other, in the *communication* treatment all participants in a group may send email messages to each other during the whole experimental session. The communication device is implemented similar to a chat forum where participants may post a message that are broadcasted to all other participants in the group.[15]

The experiment was conducted in the *Laboratory for Experimental Research* at the University of Bonn. All sessions were computerized and the software was developed by using RatImage (ABBINK and SADRIEH, 1995). In total 96 students of different disciplines were involved in the experiment – 72 take part as agents and 24 as principals. Every candidate was allowed to participate not more than in one session. We collected twelve independent observations for each treatment. After the instructions[16] participants were randomly and anonymously matched to groups of four. Additionally, one participant of each group was randomly and anonymously assigned the role of the principal. The other three took the part of the agents. The group matching was fixed for the whole experiment. A session consisted of 30 repetitions of the same tournament setting. The sessions lasted for about 1.5 to 2.5 hours. During the experiment the payoffs were given in “talers”, and in the end they were converted into Euro by a previously known exchange rate of 200 talers per 1 Euro. All subjects were paid anonymously.

---

[15] Note that this particular communication mode is prevalently used within organizations nowadays. Participants in the experiment, however, were not allowed to reveal their identity or to threaten each other or to agree upon side payments for the time after the experiment. This was ensured by monitoring the chat protocols during the experiment.

[16] A translation of the instruction sheet can be found in the appendix. Original instructions in German are available from the authors upon request.

V. Results

A. Prize Differential

Following the theoretic prediction output should be much higher if the contract provides incentives (positive wage dispersion) compared to a fixed-wage contract. Table 3 depicts average output[17] and behavior of agents in the experiment depending on wage dispersion.

Table 3: Output and behavior of agents depending on prize differential

	No Incentive	*Incentives*					Total average
	Δ_0	Average of *Incentives*	Δ_1	Δ_2	Δ_3	Δ_4	
Output	17.59	18.57	18.04	22.03	14.92	17.09	17.87
Effort	24.92	41.23	38.55	44.95	41.13	44.08	36.63
Sabotage	3.67	11.33	10.25	11.46	13.10	13.50	9.38

If one compares the output in the incentive conditions with the no-incentive condition no difference can be found at a conventional significance level.

OBSERVATION A.1: Output induced by fixed-wage contracts is similar to the output induced by incentive contracts.

Table 3 already indicates that average effort *and* average sabotage tend to increase with widening the prize differential. Compared to the theoretically predicted behavior effort and sabotage activities lie above the equilibrium level if one of the three lowest prize differentials is chosen.[18] Table 4 depicts the results of a linear regression with robust standard errors regarding the influence of the prize differential on effort and sabotage. The constant indicates the amount of activity exerted for full wage compression.

[17] In order to base our results more closely on the observed behavior in the experiment, we derive the output directly from the chosen effort and sabotage activities for our analysis without taking into account the random component. The results are qualitatively the same if one includes the random draw to determine the output. Note that the expected value of the random component is zero.

[18] By using the Binomial test the null hypothesis can be weakly rejected in favor of the alternative hypothesis that the average values per prize differential per statistically independent group are more often above the equilibrium level than below at a level of significance of at least $\alpha = .1$ (two-tailed).

Table 4 : Linear regression results – Baseline

Dependent variable:	Effort	Sabotage
Delta	0.1036*** (0.0245)	0.0500*** (0.0857)
Constant	28.0751*** (3.2806)	5.2551*** (1.1575)
	N=981 p = 0.002 R^2=0.06	N=981 p = 0.000 R^2=0.10

Note: Numbers in parentheses are robust standard errors. ***, **, * denote significance at the 1%, 5%- and 10% level, respectively.

The influence of wage compression on the amount of effort and sabotage exerted is highly significant.

OBSERVATION A.2: Effort as well as sabotage are higher the higher the prize differential.

Obviously, agents respond to strengthened tournament incentives by intensifying both activities which is qualitatively in line with standard theory. However, sabotage increases to a greater extent than the productive effort and thus, the additional effort is corroded by additional sabotage.[19]

B. Wage Sum

In the vast majority of studies on tournament incentives only the prize differential is assumed to influence agents' behavior and not the absolute level of prizes.[20] Yet, a finding from many experimental studies has been well-established in recent years that the behavior of subjects can be driven by reciprocal fairness (e.g. FEHR, GÄCHTER, and KIRCHSTEIGER 1997). In this context, a selection of the high wage level could be interpreted as a "friendly" action of the principal towards the agents, and as a consequence, agents might intend to compensate the principal for the higher costs she has to bear.[21]

[19] The ratio of sabotage to effort exerted in equilibrium results directly from the cost parameters of both activities and is 2/7. Keeping this ratio constant a rise in both activity dimensions results in an increase in output. Our finding that output is not significantly different in the incentive conditions compared to the no-incentive situation indicates that more sabotage is exerted per effort, i.e., the ratio tends to be higher in case of incentive contracts than in contracts with fixed wages.

[20] The effect of the level of prizes is empirically analyzed with data of golf tournaments in EHRENBERG and BOGNANNO (1990) and ORSZAG (1994). While EHRENBERG and BOGNANNO find that effort decreases with the prize level in later rounds of the tournament, ORSZAG argues that the prize level has no effect on the scores.

[21] For a theoretical approach to the influence of the amount of wages on effort exerted see AKERLOF and YELLEN (1990). According to their "fair wage-effort hypothesis" effort is withdrawn if the wage received is lower than the wage that is perceived to be a fair wage.

Table 5 : Impact of wage sum

	Average for low wage level W_L		Average for high wage level W_H
Output	16.07	<***	21.52
Effort	34.63	<****	43.08
Sabotage	9.28	<***	10.78

Note: **** and *** denote significance at the .5% and 1% level, respectively (Wilcoxon Signed Rank test, one-tailed).

Table 5 depicts the average results aggregated for each wage level and gives us our first result regarding the influence of the absolute wage level. An analysis of the output level shows that the effect of a higher wage sum is unambiguous:

OBSERVATION B.1*:* The output is higher for contracts with a high wage sum than for contracts with a low wage sum.

Thus, agents seem to reciprocate on an increase in wage levels by generating more output.[22] Further analysis reveals that the wage sum has an effect on both activity dimensions:

OBSERVATION B.2*:* Effort and sabotage are higher for contracts with a high wage sum than for contracts with a low wage sum.

Since the net effect of a higher wage sum on output is unambiguous, we conclude that the sabotage is less increased by a higher wage sum than effort.[23]

C. Communication

In an additional treatment our experimental design allows the principal and the three agents to communicate with each other. If communication tends to increase overall efficiency in our tournaments, agents should exert lower sabotage activities but higher effort which would result in higher output. The principal on her part should select the higher wage sum.[24] We start by analyzing the observed output.

---

[22] The principal's payoff, however, does not significantly differ between situations with different wage levels. Thus, the positive effect of higher wage on output is canceled out by the higher wage payment.

[23] The ratio of costs of sabotage to costs of effort is weakly significantly larger in the situations with the low wage level than with the high wage level (Wilcoxon Signed Rank test, $\alpha = .01$, one-tailed).

[24] Exerting sabotage decreases the payoff of both types of players. The sum of effort of all players is multiplied by a surplus factor such that the sum of costs of effort of all agents does not exceed the additional gain for the principal. Finally, the principal does not have to bear full wage costs and thus, the selection of the high wage level increases the sum of payoffs.

OBSERVATION C.1*:* In the *communication* treatment output is higher than in the *baseline* treatment.

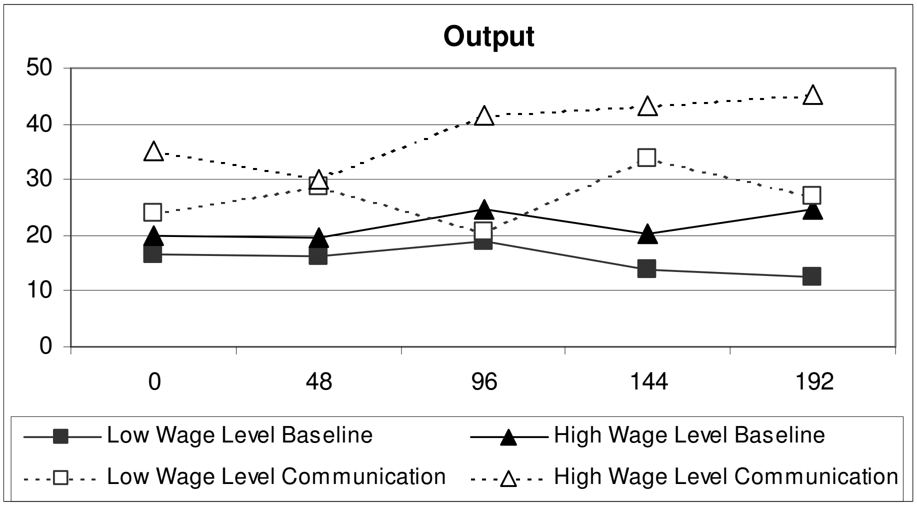

Figure 2: Average output induced by a given contract in the *baseline* and the *communication* treatment

The output in the *communication* treatment is significantly higher than in the *baseline* treatment without communication (Mann-Whitney-U test, α = .0001, one-tailed). For an illustration see Figure 2.

If one compares the activities exerted in both treatments the following result becomes evident:

OBSERVATION C.2*:* In the *communication* treatment effort is slightly higher and sabotage is substantially lower than in the *baseline* treatment

The average effort in the incentive contracts as well as in the no-incentive contracts is weakly significantly higher in the *communication* treatment than in the *baseline* (both: Mann-Whitney-U test, α = .1, one-tailed).[25] The destructive activity is even more distinctively influenced by the introduction of communication: The sabotage activity is highly significantly lower in the *communication* treatment than in the *baseline* treatment (incentive contracts:

[25] Note that this result is probably due to the frequency of the high wage level (see below) which is selected more often in the *communication* treatment and which results - according to observation B.2 - in higher effort levels. Comparing the effort levels of both treatments aggregated for each wage level yields no significant result at a conventional level.

Mann-Whitney-U test, α = .1, one-tailed and fixed-wage contracts: Mann-Whitney-U test, α = .01, one-tailed).[26]

Figure 3 shows the average effort and the average sabotage activity for a given contract. The results that are stated in subsection A. and B. are reflected by the figures and appear to be robust and in line with the behavior of agents in the *communication* treatment.[27]

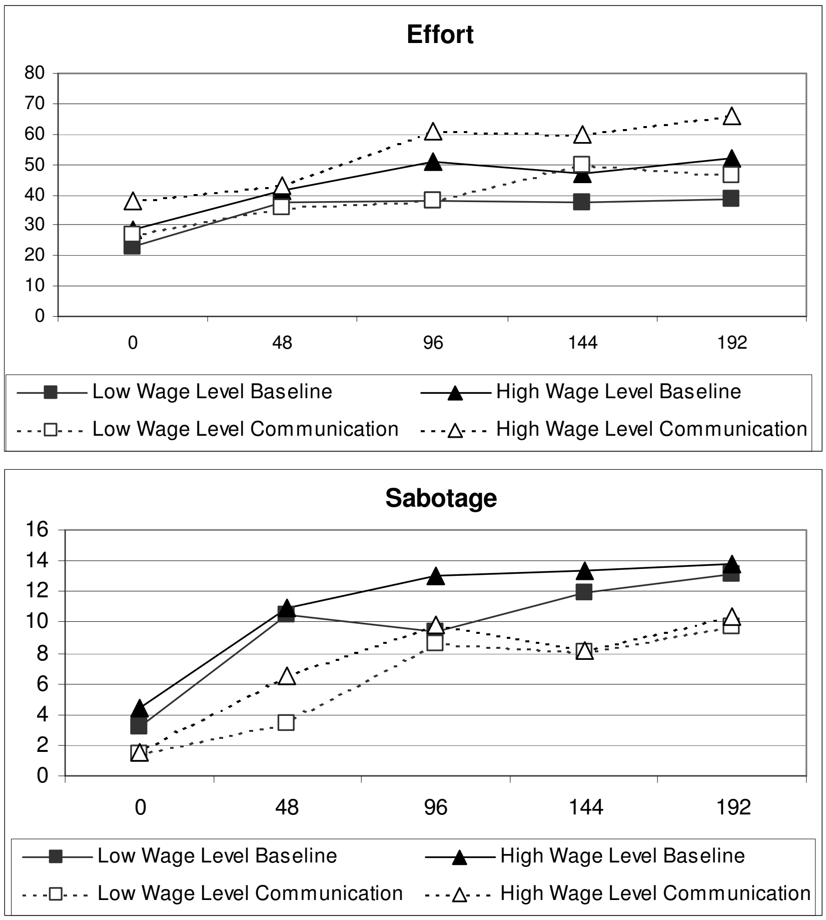

Figure 3 : Average effort and sabotage induced by a given contract in the *baseline* and the *communication* treatment

The question arises which type of contract is preferred by the principal. Figure 4 depicts the relative frequency of each wage contract in the *baseline* as well as in the *communication* treatment. Confirming our hypothesis regarding an overall increase of efficiency, we can state

[26] Comparing the wage contracts aggregated for each wage level the sabotage activity is also significantly lower with communication (low wage level: Mann-Whitney-U test, α = .05, one-tailed and high wage level: Mann-Whitney-U test, α = .01, one-tailed).

that the principal chooses contracts with the high wage level significantly more often in the *communication* treatment than in the *baseline* treatment (Mann-Whitney-U test, $\alpha = .001$, one-tailed). Moreover, fixed-wage contracts are more frequently chosen in the *communication* treatment than in the *baseline* (Mann-Whitney-U test, $\alpha = .01$, one-tailed).

OBSERVATION C.3:. In the *communication* treatment the principal selects more frequently wage contracts with the high wage level and no-incentive contracts with fixed wages than in the *baseline* treatment.

[27] The statistical analysis of the communication treatment yields the same results as described in subsections A und B.

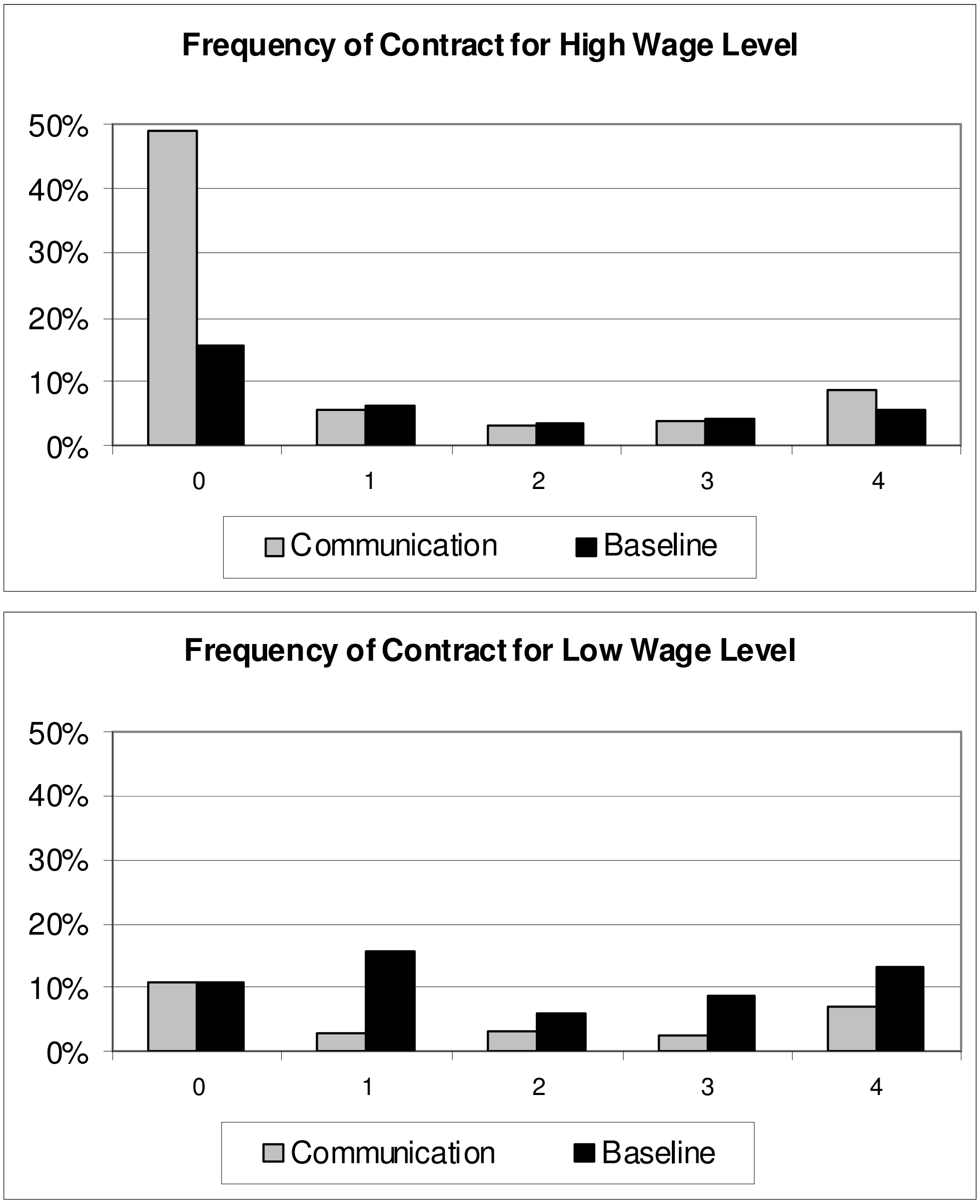

Figure 4: Relative frequency of wage contracts for low and high wage level

Figure 4 shows that the no-incentive contract implementing fixed wages with the high wage level is selected in almost half of all situations (49.17%) in the *communication* treatment and is by far the most frequently selected contract. In the *baseline* treatment no particular contract is outrageously frequently implemented by the principal although in the subgame perfect equilibrium the contract with the low wage level and the highest prize differential should be selected.

Analyzing the discussion patterns in the *communication* treatment reveals that agents in most groups ask the principal to choose the high wage level with fixed wages. Agents are offering a high effort level in return and to exert no sabotage activity.[28] The principal and agents bargain for the effort level that yields a "fair" outcome for both parties, i.e., equal payoffs for both types of players in the end.[29] Comparing the payoffs yielded by the principal with those by the agents no significant difference can be found in the *communication* treatment while agents earn significantly more than the principals in the *baseline* treatment (Wilcoxon Signed Rank test, $\alpha = .01$, one-tailed).[30] Aggregated over all contracts both types of players earn significantly more in the *communication* treatment than in the *baseline* treatment (principal: Mann-Whitney-U test, $\alpha = .001$, one-tailed and agents: Mann-Whitney-U test, $\alpha = .05$, one-tailed). In many groups the excellent "teamwork" is praised in the end, and participants express their gratefulness to the other players in their group.[31]

VI. Conclusion

Pay for individual merit that is linked to relative performance evaluation suffers from a severe drawback: agents may deteriorate the other agents' performance to improve their own relative position, i.e., they can sabotage each other. In real-world organizations data on sabotage can barely be collected as this destructive activity is strictly forbidden. Therefore, in this study the problem is tackled by an experimental approach. We focus on a situation where a principal and several agents interact: The principal commits herself to a wage contract specifying the wage level and wage dispersion, i.e., she may select fixed wages implying no incentives or a variety of competitive tournament incentives differing in the prize spread.

Standard tournament theory can be confirmed with regard to the impact of the prize differential: activities in both dimensions are intensified with augmenting the wage dispersion which intensifies competition among agents. However, sabotage increases to a greater extent.

---

[28] This agreement is quite stable in most groups although a single agent often acts as a free-rider and exerts no effort at all or only very low efforts.

[29] Approximately equal payoffs are achieved by both types of players if agents choose an effort level of 20 and no sabotage in the condition with the low wage level and an effort level of 40 and no sabotage if the high wage level is selected. In those groups that bargain for the effort level the sabotage activity is chosen to be zero in almost each round. Only in groups where communication is not focused on particular strategies of the game higher sabotage activities are exerted. For evidence of inequity aversion from experiments and an approach to model this social preferences see FEHR and SCHMIDT (1999) and BOLTON and OCKENFELS (2000).

[30] In the subgame perfect equilibrium agents earn much less than the principal.

[31] Interestingly, sabotage is also significantly reduced if we frame the instructions. In an additional *framing* treatment we replace the neutral language used in the *baseline* treatment by providing the participants with a employer-employee context. Whereas in the *baseline* treatment participants choose e.g. a "number A" and a "number B" participants in the *framing* treatment are assigned the role of an "employer" or an "employee" who chooses "effort" and "sabotage" after the "employer" has committed herself to a certain "wage contract". Thus, the context of an employer-employee relationship with the explicit term "sabotage" may lower destructive activities. Besides a lower sabotage activity all results reported from the *baseline* treatment can be confirmed for the *framing* treatment.

The obtained output reflects the result of the activities in both dimensions and is, as a consequence, not increasing with stronger incentives. Interestingly, fixed wages induce a positive output that is well above the game theoretic prediction.

If the principal commits herself to a high wage level agents increase both activities such that output increases. This observation points to a reciprocal relationship among principal and agents. Finally, if communication is allowed the principal's favorite wage contracts implement high fixed wages. Yet, both – principal and agents – earn significantly more in the communication treatment than in the *baseline* treatment and additionally yield approximately equal payoffs. Obviously, the communication device enables the players to form an implicit contract by agreeing on a certain amount of effort and almost no sabotage activity in most groups.

To conclude, our results indicate that in our context fixed-wage contracts appear to be more preferable than predicted and as least as good as incentive contracts. The reciprocal relationship is triggered by the principal's decision to offer high wages and the opportunity to communicate. Destructive activities are clearly decreased by the introduction of communication in all wage contracts. Regarding the heated debate on relative performance evaluation in organizations our study sheds some light on the question whether the fostered competition among employees is profitable for a company in all instances. If sabotage is easy this in fact might not be the case. Our results indicate that if an implicit agreement on behavior exists that can be enforced, e.g. via open communication or via some form of "corporate culture" the implementation of no-incentive contracts might be quite successful.

References

ABBINK, Klaus and Abdolkarim SADRIEH (1995) RatImage – Research Assistance Toolbox for Computer-Aided Human Behavior Experiments. *SFB Discussion Paper B-325*, University of Bonn.

AKERLOF, George A. and Janet L. YELLEN (1990) The Fair Wage-Effort Hypothesis and Unemployment. *The Quarterly Journal of Economics 105 (2)*, 255-283.

BEWLEY, Truman (1999) Why Wages Don't Fall During a Recession. Harvard University Press.

BECKER, Brian E. and Mark A. HUSELID (1992) Incentive Effects of Tournament Compensation Systems. *Administrative Science Quarterly 37*, 336-350.

BOLTON, P. and OCKENFELS, A. (2000) ERC – A Theory of Equity, Reciprocity and Competition. *American Economic Review 90,* 166-193.

BOYLE, Matthew (2001) Performance Reviews: Perilous Curves Ahead. *Fortune*, 15 May 2001.

BROSIG, J., OCKENFELS, A. and WEIMANN, J. (2003) The Effect of Communication Media on Cooperation. *German Economic Review 4 (2)*, 217-241.

BULL, Clive, Andrew SCHOTTER, and Keith WEIGELT (1987) Tournaments and Piece Rates: An Experimental Study. *Journal of Political Economy 95*, 1-33.

CAMERER, Colin (2003) Behavioral Game Theory. Princeton University Press.

CHARNESS, Gary (2000) Responsibility and effort in an experimental labor market. *Journal of Economic Behavior and Organization 42*, 375-384.

CHARNESS, G. and DUFWENBERG, M. (2002) Promises and Partnership. *Discussion Paper*, University of California and University of Stockholm.

CRAWFORD, V. (1998) A Survey of Experiments on Communication via Cheap Talk. *Journal of Economic Theory 78,* 286-298.

DEMOUGIN, Dominique and Claude FLUET (2003) Inequity Aversion in Tournaments. *Discussion Paper*, Humboldt University Berlin.

DRAGO, Robert W. and Gerald T. GARVEY (1998) Incentives for Helping on the Job. Theory and Evidence, *Journal of Labor Economics 16(1)*, 1-15.

DUFWENBERG, Martin and Georg KIRCHSTEIGER (2000) Reciprocity and Wage Undercutting. *European Economic Review 44*, 1069-1078.

EHRENBERG, Ronald G. and Michael L. BOGNANNO (1990a) The Incentive Effects of Tournaments Revisited: Evidence from the European PGA Tour. *Industrial and Labor Relations Review 43*, 74-88.

EHRENBERG, Ronald G. and Michael L. BOGNANNO (1990b) Do Tournaments Have Incentive Effects? *Journal of Political Economy 98*, 1307-24.

ERIKSSON, Tor (1999) Executive Compensation and Tournament Theory: Empirical Tests on Danish Data. *Journal of Labor Economics 17(2)*, 262-280.

FALK, Armin and Ernst FEHR (2003) Why Labour Market Experiments? *Labour Economics 10*, 399-406.

FALK, Armin and Ernst FEHR (in progress) The Power and Limits of Tournament Incentives.

FALK, Armin, Ernst FEHR, and Urs FISCHBACHER (2000) Testing Theories of Fairness – Intentions Matter. *Working Paper No. 63*, University of Zurich.

FALK, Armin and Simon GÄCHTER (2002) Reputation and Reciprocity: Consequences for the Labor Relation. *Scandinavian Journal of Economics 104 (1)*, 1-27.

FEHR, Ernst, Georg KIRCHSTEIGER, Arno RIEDL (1993) Does Fairness prevent Market Clearing? An Experimental Investigation. *Quarterly Journal of Economics 108 (2)*, 437-460.

FEHR, Ernst, Simon GÄCHTER and Georg KIRCHSTEIGER (1997) Reciprocity as a Contract Enforcement Device: Experimental Evidence. *Econometrica 64 (4)*, 833-860.

FEHR, E. and SCHMIDT, K. (1999) A Theory of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics 114,* 817-868.

FRANK, ROBERT H. and Philip COOK (1995) The Winner-Take-All Society. Free Press.

GÄCHTER, S. and FEHR, E. (1999) Collective Actions as a Social Exchange. *Journal of Economic Behavior and Organization 39,* 341-369.

GARICANO, Luis and Ignacio PALACIOS-HUERTA (2000) An Empirical Examination of Multidimensional Effort in Tournaments. Discussion Paper University of Chicago.

GIBBONS, Robert (1998) Incentives in Organizations. *Journal of Economic Perspectives 12* (Fall 1998), 115-132.

GREEN, Jerry R. and Nancy L. STOKEY (1983) A Comparison of Tournaments and Contracts. *Journal of Political Economy 91*, 349-64.

GRUND, Christian and Dirk SLIWKA (2002) Envy and Compassion in Tournaments. *IZA Discussion Paper No. 647*.

HARBRING, Christine and Bernd IRLENBUSCH (2003a) An Experimental Study on Tournament Design. *Labour Economics 10(4)*, 443-464.

HARBRING, Christine and Bernd IRLENBUSCH (2003b) Incentives in Tournaments with Endogenous Prize Selection. *mimeo*.

HARRINGTON, Joseph Jr. and Greg D. Hess (1996) A Spatial Theory of Positive and Negative Campaigning. *Games and Economic Behavior 17*, 209-29.

KAGEL, J. H. and ROTH, A. E. (1995) (eds.) Handbook of Experimental Economics. Princeton.

KONRAD, Kai (2000) Sabotage in Rent-seeking contests. *Journal of Law, Economics and Organizations 16 (1)*,. 155-165.

KRÄKEL, Matthias (2000) Relative Deprivation in Rank-Order Tournaments. *Labour Economics 7*, 385-407.

LAZEAR, Edward P. (1989) Pay Equality and Industrial Politics. *Journal of Political Economy 97*, 561-80.

LAZEAR, Edward P. (1999) Personnel Economics: Past Lessons and Future Directions – Presidential Address to the Society of Labor Economists. *Journal of Labor Economics 17(2)*, 199-236.

LAZEAR, Edward P. and Sherwin H. ROSEN (1981) Rank-Order Tournaments as Optimum Labor Contracts. *Journal of Political Economy 89*, 841-64.

LYNCH, Jim and Jeffrey S. ZAX (1998) Prizes, Selection and Performance in Arabian Horse Racing. *Discussion Paper,* University of Colorado.

MAIN, Brian G.M., Charles A. O'REILLY III, and James WADE (1993) Top Executive Pay: Tournament or Teamwork? *Journal of Labor Economics 11(4)*, 606-28.

MOOD, Alexander M., Franklin A. GRAYBILL, and Duane C. BOES (1974) Introduction to the Theory of Statistics. 3rd Edition. McGraw-Hill.

MURPHY, Kevin J. (1992) Performance Measurement and Appraisal: Motivating Managers to Identify and Reward Performance. William J. Bruns Jr. (ed.) Performance Measurement, Evaluation, and Incentives. Harvard Business School Press.

NALBANTIAN, Haig R. and Andrew SCHOTTER (1997) Productivity under Group Incentives: An Experimental Study. *American Economic Review 87*, 314-41.

NALEBUFF, Barry J. and Joseph E. STIGLITZ (1983) Prizes and Incentives: Towards a General Theory of Compensation and Competition. *Bell Journal of Economics 14*, 21-43.

O'KEEFFE, Mary, W. Kip VISCUSI, and Richard J. ZECKHAUSER (1984) Economic Contests: Comparing Reward Schemes. *Journal of Labor Economics 2* (January 1984), 27-56.

ORBELL, J. M., DAWES, R. M. and VAN DE KRAGT, A. J. C. (1988) Explaining Discussion-Induced Cooperation. *Journal of Personality and Social Psychology 54,* 811-819.

ORRISON, Alannah, Andrew SCHOTTER, and Keith WEIGELT (1997) On The Design of Optimal Organizations Using Tournaments: An Experimental Examination. *Discussion Paper*, New York University.

ORSZAG, Jonathan M. (1994) A New Look at Incentive Effects and Golf Tournaments. *Economics Letters 46*, 77-88.

OSTROM, E. and WALKER, J.M. (1991) Communication in a Commons: Cooperation without External Enforcement. *Laboratory Research in Political Economy*, 287-322.

PFEFFER, Jeffrey and Robert I. SUTTON (1999). The Knowing-Doing Gap: How Smart Companies Turn Knowledge into Action. Harvard Business School Press, Cambridge.

PRENDERGAST, Canice (1999) The Provision of Incentives in Firms. *Journal of Economic Literature 37*, 7-63.

ROSEN, Sherwin H. (1986) Prizes and Incentives in Elimination Tournaments. *American Economic Review 76*, 701-15.

SALLY, D. (1995) Conversation and Cooperation in Social Dilemmas: a Meta-Analysis of Experiments from 1958 to 1992. *Rationality and Society (7)*, 58-92.

SCHOTTER, Andrew and Keith WEIGELT (1992) Asymmetric Tournaments, Equal Opportunity Laws, and Affirmative Action: Some Experimental Results. *Quarterly Journal of Economics 107*, 511-39.

SKAPERDAS, Stergios and Bernard GROFMAN (1995) Modeling Negative Campaigning. *American Political Science Review 89*, 49-61.

VAN DIJK, Frans, Joep SONNEMANS, Frans VAN WINDEN (2001) Incentive Systems in a Real Effort Experiment. *European Economic Review 45*, 187-214.

WEIGELT, Keith, Janet DUKERICH and Andrew SCHOTTER (1989) Reactions to Discrimination in an Incentive Pay Compensation Scheme: A Game-Theoretic Approach. *Organizational Behavior and Human Decision Processes 44*, 26-44.

Appendix: Marginal Probabilities of Winning

We show that given our assumptions in a symmetric equilibrium the marginal probabilities of winning depend only on the size of the interval from which the random component in the production function is drawn. In what follows we concentrate on the marginal probability regarding the chosen effort level. Let λn be the number of winner prizes ($0 < \lambda < 1$), i.e., agents with the $(1 - \lambda)n$th lowest output or less receive the loser prize. In a symmetric Nash equilibrium each competitor of agent i will choose the same effort level e^* and the same sabotage level s^*. Therefore, agent i will receive the winner prize if her output is higher than the $(1 - \lambda)n$th lowest output of the other $(n - 1)$ workers, i.e., if $e_i + \varepsilon_i - (n - 1)\, s^* > e^* - s_i - (n - 2)\, s^* + \hat{\varepsilon}$ with $\hat{\varepsilon}$ as the $(1 - \lambda)n$th lowest of $(n - 1)$ order statistics.

The probability for this event is $\Pr\{X < e_i - e^* - (n - 1)\, s^* + s_i + (n - 2)\, s^*\} = F_X(e_i - e^* + s_i - s^*)$ with $X := \hat{\varepsilon} - \varepsilon_i$ and $F_X(\cdot)$ as the distribution function of X. Agent i maximizes her expected payoff $E\Pi_i(e_i) = m + \Delta\, F_X(e_i - e^* + s_i - s^*) - C_e(e_i) - C_s(s_i)$. From the assumption of a symmetric equilibrium ($e_j = e^*$ and $s_j = s^*$ for $j = 1, \ldots, n$) it follows that the equilibrium effort is characterized by $C_e'(e+) = \Delta f_X(0)$ and $C_s'(s+) = \Delta f_X(0)$ with $f_X(\cdot) = F'_X(\cdot)$ as X's density function. This leads to $e^* = C_e'^{-1}(\Delta f_X(0))$ as well as $s^* = C_s'^{-1}(\Delta f_X(0))$ with $C'^{-1}(\cdot)$ as the inverse function of the marginal cost function (note that $C_e'^{-1}(\cdot)$ and $C_s'^{-1}(\cdot)$ are linearly increasing). In order to obtain the equilibrium effort it remains to derive the explicit probability $f_X(0)$.

Let $F(\varepsilon_j)$ and $f(\varepsilon_j)$ be the distribution function and the density function of each of the $j = 1, \ldots, n$ i.i.d. random components ε_j. The density function of the $(1 - \lambda)n$th lowest of $(n - 1)$ order statistics can be written as[32]

$$
\begin{aligned}
f_{(1-\lambda)n;(n-1)}(\hat{\varepsilon}) &= \frac{(n-1)!}{[(1-\lambda)n-1]!\,[\lambda n-1]!} F^{\,n-\lambda n-1}(\hat{\varepsilon})\,[1-F(\hat{\varepsilon})]^{\lambda n-1} f(\hat{\varepsilon}) \\
&= \frac{(n-1)!}{[(1-\lambda)n-1]!\,[\lambda n-1]!}\left[\frac{\bar{\varepsilon}+\hat{\varepsilon}}{2\bar{\varepsilon}}\right]^{n-\lambda n-1}\left[1-\frac{\bar{\varepsilon}+\hat{\varepsilon}}{2\bar{\varepsilon}}\right]^{\lambda n-1}\frac{1}{2\bar{\varepsilon}} \\
&= \frac{(n-1)!}{[(1-\lambda)n-1]!\,[\lambda n-1]!}\,\frac{(\bar{\varepsilon}+\hat{\varepsilon})^{n-\lambda n-1}(\bar{\varepsilon}-\hat{\varepsilon})^{\lambda n-1}}{(2\bar{\varepsilon})^{n-1}}
\end{aligned}
$$

Now let us consider the density function $f_X(x)$ for the random variable $X := \hat{\varepsilon} - \varepsilon_i$. Because ε_i and $\hat{\varepsilon}$ are stochastically independent, we have

$$
f_X(x) = \int_{X=\hat{\varepsilon}-\varepsilon_i} f(\varepsilon)\, f_{(1-\lambda)n;(n-1)}(\varepsilon + x)\, d\varepsilon
$$

$$= \frac{(n-1)!}{[(1-\lambda)n-1]!\,[\lambda n-1]!} \int_{X=\hat{\varepsilon}-\varepsilon_i} \frac{(\bar{\varepsilon}+\varepsilon+x)^{n-\lambda n-1}(\bar{\varepsilon}-\varepsilon-x)^{\lambda n-1}}{(2\bar{\varepsilon})^n}\, d\varepsilon$$

In order to compute the density function $f_X(x)$ we have to fill in the limits of the integral. We know that

$-\bar{\varepsilon} \le \varepsilon \le \bar{\varepsilon}$ (*)

$-\bar{\varepsilon} \le \hat{\varepsilon} \le \bar{\varepsilon} \Leftrightarrow -\bar{\varepsilon} \le \varepsilon + x \le \bar{\varepsilon} \Leftrightarrow -x-\bar{\varepsilon} \le \varepsilon \le \bar{\varepsilon} - x$. (**)

The random variable $X := \hat{\varepsilon} - \varepsilon_i$ is distributed over the interval $[-2\bar{\varepsilon}, 2\bar{\varepsilon}]$, which can be divided into two subintervals: (i) $[-2\bar{\varepsilon}, 0]$ and (ii) $[0, 2\bar{\varepsilon}]$. For $x \le 0$ we obtain from (i) together with (*) and (**) that $-x-\bar{\varepsilon} \le \varepsilon \le \bar{\varepsilon}$. Accordingly, for $x > 0$ we obtain from (ii) together with (*) and (**) that $-\bar{\varepsilon} \le \varepsilon \le \bar{\varepsilon} - x$.

This gives us

$$f_X(x) = \begin{cases} \dfrac{(n-1)!}{[(1-\lambda)n-1]!\,[\lambda n-1]!} \displaystyle\int_{-x-\bar{\varepsilon}}^{\bar{\varepsilon}} \frac{(\bar{\varepsilon}+\varepsilon+x)^{n-\lambda n-1}(\bar{\varepsilon}-\varepsilon-x)^{\lambda n-1}}{(2\bar{\varepsilon})^n} d\varepsilon & \text{if } -2\bar{\varepsilon} \le x \le 0 \\[2ex] \dfrac{(n-1)!}{[(1-\lambda)n-1]!\,[\lambda n-1]!} \displaystyle\int_{-\bar{\varepsilon}}^{\bar{\varepsilon}-x} \frac{(\bar{\varepsilon}+\varepsilon+x)^{n-\lambda n-1}(\bar{\varepsilon}-\varepsilon-x)^{\lambda n-1}}{(2\bar{\varepsilon})^n} d\varepsilon & \text{if } 0 < x \le 2\bar{\varepsilon} \end{cases}$$

In equilibrium it holds that $x = 0$. It follows that

$$f_X(0) = \frac{(n-1)!}{[(1-\lambda)n-1]!\,[\lambda n-1]!\,(2\bar{\varepsilon})^n} \int_{-\bar{\varepsilon}}^{\bar{\varepsilon}} (\bar{\varepsilon}+\varepsilon)^{n-\lambda n-1}(\bar{\varepsilon}-\varepsilon)^{\lambda n-1} d\varepsilon$$

Repeated partial integration gives

$$\int_{-\bar{\varepsilon}}^{\bar{\varepsilon}} (\bar{\varepsilon}+\varepsilon)^{n-\lambda n-1}(\bar{\varepsilon}-\varepsilon)^{\lambda n-1} d\varepsilon =$$

$$= \left[\frac{1}{n-\lambda n}(\bar{\varepsilon}-\varepsilon)^{\lambda n-1}(\bar{\varepsilon}+\varepsilon)^{n-\lambda n} + \frac{(\lambda n-1)}{(n-\lambda n)(n-\lambda n+1)}(\bar{\varepsilon}-\varepsilon)^{\lambda n-2}(\bar{\varepsilon}+\varepsilon)^{n-\lambda n+1} + \ldots + \right.$$

$$\left. + \ldots + \frac{(\lambda n-1)!}{(n-\lambda n)\ldots(n-1)}(\bar{\varepsilon}+\varepsilon)^{n-1} \right]_{-\bar{\varepsilon}}^{\bar{\varepsilon}}$$

$$= \left[\frac{(\lambda n-1)!}{(n-\lambda n)(n-\lambda n+1)\ldots(n-1)}(2\bar{\varepsilon})^{n-1} \right]$$

Thus, the density reduces to

$$f_X(0) = \frac{1}{2\bar{\varepsilon}}$$

[32] For the construction of order statistics see MOOD, GRAYBILL and BOES (1974).

Appendix: Instructions

{*baseline* treatment, original instructions were in German; they are available from the authors upon request}

Rounds, Groups and Roles

- You are participating in an experiment of **30 rounds.**
- You will be assigned to a group of 4 participants. During the experiment you will only interact with participants of your group. The group is randomly assigned and is kept constant throughout the whole experiment. It is not announced which participants are assigned to one group. There are two red and two blue groups.
- Each participant has one of two roles: **1 participant** is of **type I** and **3 participants** are of **type II**.
- Costs and payoffs are given in the fictitious currency „taler".

Procedure at the first stage – Decision of participant of type I

- The participant of type I decides, whether he wants to offer a payment to participants of type II or not.
- If he does not offer a payment all players including player of type I receive a payment of 0 taler in this round, and the next round begins.
- If he offers a payment he must decide whether the total sum of payments which participants of type II are going to receive amounts to 300 or 600 taler.
- Moreover, he selects one of 5 feasible distributions of the total sum of payments.
- The participant of type I may decided that all participants of type II receive the same payment or that two participants of type II obtain a low payment and one receives a high payment. The amount of the low and high payment must be selected as well.

total sum of payments: **300 taler**		total sum of payments: **600 taler**	
100 taler for each participant of type II		**200 taler for each participant of type II**	
low payment	**high payment**	**low payment**	**high payment**
84	132	184	232
68	164	168	264
52	196	152	296
36	228	136	328

- The task of the participants of type II is to choose a number A and a number B in each round which determine the result.

Procedure at the second stage – Decision of participants of type II

- The participants of type II may take a decision if the participant of type I has offered a payment to them at the first stage.
- The participants of type II are informed about the amount of the total sum of payments and whether all payments are equal or what the amounts of the low and the high payments are.
- Each participant of type II chooses a **number A** from the set {0, ..., 100} and a **number B** from the set {0, ..., 50}. For both numbers a participant has to bear the costs listed in the two cost tables. The higher the numbers chosen the higher are the costs.
- For each participant of type II a random number is drawn independently from the set {–60, ..., +60}. Each number is drawn with the same probability

Payoff of participants of type II:

- If the participant of type I has decided at the first stage that all participants of type II shall receive the same payment, each participant of type II receives the same selected payment.
- If the participant of type I has decided that there was one high payment and two low payments, the value of the results of the participants of type II determine the amount of the payoff each player is going to receive. The participant with the highest result receives a high payment, and the two participants with the lowest results receive a low payment. (In case of identical results a fair random move decides who receives a high and who a low payment.)
- The **result** of a participant is the sum of her/his random number and the number A chosen by her/him. Moreover, all numbers B are deducted which the two *other* participants of type II have chosen. Thus, the result of a participant of type II increases with the own number A and decreases with the numbers B of the other three participants of type II.

result = own number A - numbers B of other participants + random number

- The cost for number A and for number B – which the participant has chosen– are subtracted from this payment. This results is the **payoff of the round**

round payoff of participant of type II= payment – costs for number A and number B

- After each round the participant of type II is informed about his own payoff, round payoff, the results of all type II participants and the round payoff of participant of type I.

Payoff of participant of type I:

- The round payoff of the participants of type I is as follows:

round payoff of participant of type I= 3*(sum of results of type II participants) – 0,3*(total sum of payments for type II participants)

- This means: for a total sum of payments of 300 taler costs of 0.3 * 300 = 90 taler are deducted from the type I participant's payoff and for a total sum of payments of 600 taler costs of 0.3 * 600 = 180 taler are deducted.
- After each round the participant of type I is informed about the results and the payments that each of the three participants of type II received (but not about the round payoffs).

In the beginning of the experiment all participants receive a lump sum of 1200 taler.

At the end of the experiment the sum of all round payoffs is exchanged at an exchange rate of 1 Euro per 200 taler.