The Power and Limits of Tournament Incentives

Armin Falk[1], Ernst Fehr[2], David Huffman[3]

March 28th, 2008

Abstract

Tournaments provide incentives through the prize spread. Agents are predicted to work harder for higher prize spreads, and thus principals are predicted to maximize the spread. This paper shows experimentally how changing institutional environments affect the way that principals structure tournament incentives, and the degree of wage compression. While in some settings tournaments provide powerful incentives, and principals maximize the prize spread, two specific factors – sabotage opportunities, and loss aversion among agents – are shown to undermine the power of tournaments and cause principals to choose wage compression.

Keywords: Relative Pay, Wage Compression, Sabotage, Loss Aversion, Experiment

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

[1] University of Bonn and CEPR; e-mail: armin.falk@uni-bonn.de
[2] University of Zurich; e-mail: fehr@iew.uniz.ch
[3] Swarthmore College; e-mail: dhuffma1@swarthmore.edu

1 Introduction

Rank order tournaments are a powerful incentive device. By awarding high and low prizes based on relative performance, a principal can achieve stronger incentives than with a scheme that involves the same wage bill but equal wages. Theory predicts that in a tournament setting only the difference between the high and the low prize matters, rather than the absolute payment levels. An agent's performance is increasing in the spread between the winner and loser prize, ceteris paribus, and thus a principal maximizes incentives by maximizing the prize spread (see, e.g., Lazear and Rosen, 1981).

Although tournaments are a prevalent form of compensation, for example arising when promotions are used to provide incentives, in practice principals sometimes seem to prefer substantial wage compression relative to a policy of maximizing the prize spread (Lazear, 1989), and in other cases principals choose to avoid tournaments altogether, paying equal wages regardless of relative performance (Nalebuff and Stiglitz, 1983). The factors that cause principals to adjust the structure of tournament incentives, and choose the degree of wage compression, are not fully understood. This partly reflects a lack of empirical evidence on how principals adjust the structure of tournament incentives depending on the institutional environment.

This paper presents results from experiments that demonstrate the power of tournament incentives, but also some important limitations depending on the institutional setting, which help explain why principals may sometimes choose wage compression. Limitations arise due to features of the environment that cause agents to be unresponsive to high wage spreads. This in turn makes it optimal for principals to compress wages rather than selecting the largest possible prize spread.

Our focus is on how principals choose the structure of tournament incentives, so our experiment implements a setting where tournament incentives are *endogenous*. Principals choose from a menu of contracts differing only in terms of the prize spread. This aspect of the experiment makes it possible to study how principals choose contracts, and the degree of wage compression, depending on the work setting, and how this depends in turn on the ways that agents respond to incentives in different settings.

The power of tournaments with large prize spreads is shown in a baseline treatment where a principal is matched each period with two agents. The principal chooses a contract from a menu of contracts, which in terms of the wage spread but hold the wage sum constant. After being informed about the contract, agents choose their effort levels. The payoff to the principal depends on the outputs of the two agents, plus random shocks to productivity. The payoff to agents depends on their relative performance, and the wage spread specified in the contract.

The results from the baseline treatment provide a strong demonstration of the power of tournament incentives. Principals choose primarily the contracts with the highest wage spreads, leading to substantial wage dispersion. Agents' efforts are monotonically increasing in the prize spread, holding the wage sum constant. Effort levels are also remarkably close to the theoretical predictions for each spread. Given the behavior of agents, profits for principals are strongly increasing in the wage spread, which makes the wage setting policies of principals quite rational.

We then turn to two key factors that may undermine the effectiveness of tournament incentives, by decreasing the responsiveness of agents to high prize spreads. One is the opportunity for sabotage. As shown theoretically by Lazear (1989), if the work environment allows opportunities for agents to sabotage each other, then increasing the prize spread can be counterproductive, causing agents to shift effort from performance to sabotaging opponents in the tournaments. As a result, it can be optimal for principals to choose smaller wage spreads.

We implement a second treatment, where we added the possibility for agents to engage in a costly sabotage activity, which reduces the other agents output to zero. Given the parameters of the experiment, the theoretical prediction involves agents choosing to sabotage, and put in minimal effort, once the prize spread exceeds a critical threshold.

Sabotage opportunity dramatically changes principals' contract choices. As predicted, principals tend to choose lower wage spreads than in baseline, leading to substantial wage compression. The effort levels of agents are increasing at first in the wage spread, but then drop to the minimum level for higher spreads as workers substitute from effort to sabotage, which is also qualitatively in line with the theory. Interestingly, although overall the effort and sabotage profile is quite close to the predictions, the first sharp drop in effort, and increase in sabotage, begins, at a lower wage spread than predicted. Turning to the outcomes for principals, we find that principals' profits are decreasing in the wage spread for spreads above the critical threshold, in contrast to the baseline where the highest wage spreads were best. This means that wage compression pays, and explains the different wage setting policies of principals in the presence of sabotage opportunities.

Another factor that potentially limits the effectiveness of tournaments is *loss aversion* (Kahneman and Tversky, 1979). A large body of evidence shows that people tend to evaluate outcomes in relative terms, comparing to a reference outcome, and that they are particularly averse to losses relative to this reference level. The effectiveness of tournament incentives may depend on whether agents view the outcome of working hard and then getting the loser prize as simply a smaller gain, or as truly being a loss relative to a reference level. Given that a firm is limited in terms of the wage bill it can pay, increasing the spread necessarily means lowering the loser prize, and thus larger spreads may be more likely to drive the loser payoff below an agents' reference payoff. Loss aversion on the part of agents could thus limit their willingness to work hard as the prize spread increases: while working hard makes success more likely, the extra cost makes the outcome of losing the tournament even worse. As a result, large prize spreads could work less well, and principals could find it optimal to choose some wage compression, in settings where agents face the potential for losses.

The loss aversion treatment keeps incentives exactly the same as in baseline, because the wage spreads for different contracts are unchanged. The only difference is that the constraint on the principal's wage sum is varied: the principal has to pay a lower wage sum than in baseline. The wage sum is sufficiently low that for high wage spreads, the outcome of working hard and then getting the low prize falls below what is likely to be a salient reference point for agents, namely zero. For low wage spreads there is no possibility of falling below zero, but for high wage spreads working hard and then losing the tournament leads to a negative payoff. Note that agents participate in multiple tournaments, and there is no possibility for agents to have negative total earnings for the experiment due to a show up fee. Nevertheless, agents may dislike falling below zero in any given tournament interaction. The standard theoretical prediction is identical to that for the baseline treatment, because incentives in terms of the spread are unchanged. If agents are loss averse, however, then the change in the wage sum could lead to wage compression, because loss averse workers are less responsive to high wage spreads and principals adopt compressed wage structures accordingly.

In the loss treatment, principals do in fact choose more compressed wages than in the baseline, correctly anticipating loss averse behavior by agents. Agents respond to low spreads in a similar way to in baseline, but become significantly less responsive to the wage spread than in baseline at exactly the point where the spread becomes large enough that choosing a high effort level and getting the loser prize would entail a payoff negative payoff for that tournament interaction. Due to the loss averse behavior by agents, principals' profits are increasing monotonically for low prize spreads, and level off abruptly once spreads are high enough to impose losses, helping explain why the largest prize spreads are no longer predominant and wage compress is observed relative to baseline.

In summary, the findings of the paper strongly support some of the central predictions of tournament theory, while revealing some nuances of behavior, such as excess sabotage, and agent loss aversion, that are not fully understood. They also show how principals' choices of tournament structure are affected by sabotage opportunities and the potential for losses, and suggest that both sabotage and loss aversion are candidates for explaining wage compression.

Our research complements the previous empirical literature on tournament incentives. For example, Ehrenberg and Bognanno (1990) use golf data, and find that performance is increasing in the prize spread, consistent with theory.[1] A key challenge when using field data, however, is that the form of utility function, information conditions, etc. are not observed, which limits the types of predictions that can be tested. Studying sabotage is also very difficult in the field, given that by its nature sabotage is typically covert and unobservable. The role of loss aversion is also difficult to study, because assessing the potential for losses may depend on knowing agents' cost functions for effort. For these reasons an experiment is a useful tool for testing tournament theory.

Furthermore, as noted by Prendergast (1999), while empirical studies have provided useful tests of incentives, there has been less evidence on whether principals design contracts with these incentives in mind.[2] In other words, while it has been shown that increasing prize spreads can enhance agent performance, there is less evidence on why and in what circumstances principals choose varying degrees of tournament incentives and wage compression. In our experiment we can study principals' choices of incentives, knowing the menu of feasible contracts available, and can provide clear evidence on how principals' choices respond to exogenously varying environments. Of previous experiments on tournaments (e.g., Bull et al., 1987; Schotter and Weigelt, 1992; Harbring and Irlenbusch, forthcoming), ours is one of the few to give principals' control over the incentive structure, one exception being Harbring and Irlenbusch (2005). In their study, however, the focus is on how agents' behaviors are influenced by whether the tournament incentives were chosen by a principal or by the experimenter. They do not study how principals' choices respond to different institutional settings, as we do by varying the presence or absence of sabotage, or the potential for losses.[3] To our knowledge ours is the first experiment to study how varying the potential for losses affects behavior of

[1] Becker and Huselid (1992) find similar results for NASCAR drivers, and Knoeber and Thurman (1994) in the market for chicken broilers.

[2] One exception is evidence from data on CEO compensation showing that the winning prize increases in the number of participants in the tournament. This is consistent with predictions regarding the optimal tournament design by principals (see, e.g., Eriksson, 1999).

[3] Other differences include the fact that the principal in Harbring and Irlenbusch (2005) interacts with the same agents repeatedly, and the wage sum is varied across contracts simultaneously with the wage spread.

agents and principals in a principal-agent framework.

The rest of the paper is organized as follows. Section 2 explains the design of the experiment, and Section 3 derives the predictions for behavior. Section 4 presents the empirical findings and Section 5 concludes.

2 Design

2.1 Basic Design

In all treatment conditions one principal is in each period randomly matched with two agents. The principal and the agents then play a two-stage game. In the first stage the principal selects a contract, which defines the tournament incentives. He can choose from a set of eight different contract types, which differ in the wage specified for the agent with the higher output, w_h, and the wage for the agent with the lower output, w_l. The possible contracts are shown in Table 1. For example, if the principal chooses contract type 1, both agents receive the same wage (150) regardless of their output, i.e., incentives in the form of the wage spread, Δ, are zero. Contract types 2 to 8 imply an increasing strength of the incentives ranging from 20 to 140. The principal is free to choose any contract he wants to. Note that the wage sum of w_l and w_h is 300 for all contract types.

[Table 1 about here]

The choice of the contract type is transmitted to the two agents a principal is matched with. In the second stage, the agents have now to decide on a costly effort level, e, with $e \in [1, 1.5, 2, ..., 12]$. Effort costs $c(e)$ are given by $c(e) = (1 - e)^2$. It follows that for the lowest effort choice, effort costs are zero while higher effort levels are increasingly costly. Agents make their effort decisions simultaneously, i.e., without knowing the other agent's effort decision.

The output y_i of an agent i depends on his effort choice, e_i, and a random variable z_i which is drawn from the uniform distribution $[0, 10]$. Output is given by:

$y_i = e_i + z_i$. The variable z is for all agents randomly and independently determined by the computer. It reflects the fact that effort is not verifiable and that the principal can only contract on y and not on e. After agents have made their effort decisions and the computer has randomly determined z_i for all agents, payoffs are determined. A principal's payoff π_i is given according to the following formula:

$$\pi_i = \alpha y_1 + \alpha y_2 \tag{1}$$

where α was set equal to 8 in the experiment. The payoff of an agent U_i is calculated as:

$$U_i = \begin{cases} w_h - c(e_i) & \text{if } y_i > y_j \\ w_l - c(e_i) & \text{if } y_i < y_j \\ 0.5w_h + 0.5w_l - c(e_i) & \text{if } y_i = y_j \end{cases} \tag{2}$$

where w_h and w_l are as specified in the contract chosen by the principal.

At the end of the second stage each principal is informed about the outputs of both agents and the resulting payoff. Each agent is informed about his output and his payoff. Principals know their own payoff function as well as the payoff function of the agents. This is necessary for them to calculate the optimal effort decision given a contract, and thus the optimal contract. Agents on the other hand are only informed about their own payoff function. They do not know their principal's payoff function. All they know is that a principal's payoff increases in the output and thus in the chosen effort. The reason why we did not inform agents about the principal's payoff function was to limit possible fairness considerations. Previous experiments have shown that if people can calculate each other's payoffs, fairness concerns may strongly influence behavior in principal agent experiments (e.g., Fehr and Falk, 1999). Since we are in this study interested in how loss aversion and sabotage affects efforts and wages, we want to exclude possible confounds due to fairness considerations.

To study convergence of contract and effort choices, we implemented the common method of stationary replication, i.e., each treatment lasted for 12 periods. In each session we had 24 subjects. Before a session started subjects were randomly divided into 8 principals and 16 agents. They were given their instructions and had to answer several control questions to check their understanding of the experimental procedures. During the experiment principals and agents were randomly matched into groups of three, one principal and two agents. Subjects were informed about this procedure. All interaction was completely anonymous and subjects did not learn the identity of whom they were interacting with. The experiment was programmed and conducted with the software z-tree (Fischbacher, 2007). During the experiment subjects could earn Guilders, in addition to their show up fee. Payments were transformed into CHF at the end of the experiment at a rate of 100 Guilders = 0.12 CHF (1 CHF $\sim$ 0.90 US dollars). On average subjects earned about 13.42 CHF for participating in the experiment.

2.2 Treatments

We study three treatments. In the Baseline-treatment (henceforth B-treatment), principals select a contract according to Table 1, and agents choose their effort where the set of feasible efforts and associated effort costs are as outlined above. The B-treatment informs us about how agents respond to different tournament incentives and given their effort behavior how principals select contracts. Principals can either choose high powered or low powered incentives.

We compare the effort choices and the wage setting of principals in the B-treatment with those of a treatment where agents can engage in sabotage activity. In this Sabotage-treatment (henceforth S-treatment), everything is exactly as in the B-treatment except that agents now have to make two choices. First, they decide on an effort level just as in the B-treatment (with the same feasible effort levels and the same effort costs). Second, they have to decide whether they want to sabotage the other agent's output.[4] Sabotage is a binary choice. If an agent i decides to sabotage agent j's output y_j is set to zero independent of e_j and z_j. If i decides not to sabotage, y_j is determined exactly as in the B-treatment. To sabotage is

[4] In the instructions (see appendix) we avoided any value-laden term. Instead of talking about sabotage we used the phrase "you can set the other agents output to zero".

costly. If agent i sabotages the output of agent j, i incurs costs of 27 guilders. Both decisions, the effort choice and the sabotage decision are made on a computer screen presented to the agents. Both agents make their decisions simultaneously, i.e., without knowing how the other agent decides (on effort and sabotage). After the second stage of the S-treatment, principals are informed about the outputs and their payoffs, and each agent is informed about his output, whether his output was subject to sabotage, and his payoff.

Finally, we compare the B-treatment with the Loss-treatment (henceforth L-treatment). The L-treatment is exactly the same as the B-treatment with the only exception that the wage sum for each contract is no longer 300 as shown in Table 1 but 140. This does not, however, change the strength of the incentives (Δ), which range from 0 to 140: In the L-treatment the first contract (w_h, w_l) is (70, 70) with $\Delta = 0$, the second is (80, 60) with $\Delta = 20$ and so on up to the eighth contract which implies (140, 0) and $\Delta = 140$. Since the optimal effort is a function of Δ, the theoretical predictions for both the B- and the L-treatment are the same. Different from the B-treatment, however, in the L-treatment it is possible for agents to incur losses (see our discussion below).

We employed the following sequence of treatments. There were five sessions where subjects first participated in the B-treatment (B1) and then in the S-treatment. In another five sessions subjects first participated in the L-treatment and then in the B-treatment (B2). Each session was conducted with different subjects. Subjects were informed about the second treatment only after they had concluded the first treatment. Each treatment lasted for 12 rounds. The first five sessions allow a comparison between the B- and the S-treatment while we conducted the second five sessions to study possible loss-aversion effects. Moreover, a comparison between the two B-treatments, B1 and B2, informs us about the robustness of our findings in the Baseline-treatment and possible order or learning effects.

3 Predictions

The theoretical prediction for a tournament with one principal and two agents i and j is straightforward (see, e.g., Lazear and Rosen, 1981, who were the first to analyze this type of game).[5] Throughout this section we assume that subjects are risk neutral. Using the variables introduced above, and fixing a particular contract type the optimal effort choice for agent i maximizes his expected wage, net of effort costs. The effort pair e_i^* and e_j^* is a Nash equilibrium if for all i, e_i solves

$$\max_{e_i} w_h prob\{y_i(e_i) > y_j(e_j^*)\} + w_l prob\{y_i(e_i) \leq y_j(e_j^*)\} - c(e_i) \tag{3}$$

where $y_i = e_i + z_i$ and z_i and z_j are drawn independently from a density $f(z)$. The first order condition for (3) yields

$$(w_h - w_l)\frac{\partial prob\{y_i(e_i) > y_j(e_j^*)\}}{\partial e_i} = c'(e_i) \tag{4}$$

where $prob\{y_i(e_i) > y_j(e_j^*)\} = \int_{z_j} \left[1 - F(e_j^* - e_i + z_j)\right] f(z_j)dz_j$. Putting this expression into (4), and differentiating (4) yields

$$(w_h - w_l)\int_{z_j} f(e_j^* - e_i + z_j)f(z_j)dz_j = c'(e_i) \tag{5}$$

In a symmetric Nash equilibrium, (5) can be written as:

$$(w_h - w_l)\int_{z_j} (z_j)^2 dz_j = c'(e^*) \tag{6}$$

Given the parameters in the experiment and given that z is drawn from a uniform distribution on $[0, 10]$ the expression $\int_{z_j}(z_j)^2 dz_j$ amounts to $\int_0^{10}(\frac{1}{10})^2 dz_j = \frac{1}{10}$. Thus

$$(w_h - w_l)\frac{1}{10} = c'(e^*) \tag{7}$$

We can use (7) in combination with $c(e) = (1 - e)^2$ to calculate the (symmetric) equilibrium effort levels for each of the eight possible contracts. These predictions are

[5] For the following derivation of the equilibrium see Gibbons (1992).

displayed in Table 2. Since the optimal effort decision depends only on the difference between the high and the low wages and not on their levels, the predictions for the B- and the L-treatment are the same.[6]

Table 2 reveals that e^* monotonically increases in Δ. Since a principal's profit increases in the effort choices of the agents (note that $\frac{\partial \pi_i}{\partial e_i} = \frac{\partial \pi_i}{\partial e_j} = 1 > 0$), the optimal contract maximizes Δ, i.e., each principal should on the first stage choose contract type 8 with $\Delta^* = 140$. This holds in both the B- and the L-treatment.

[Table 2 about here]

Unlike the B-treatment, agents in the S-treatment have two possibilities to create a higher output than their competitor, to work hard or to sabotage. The predictions for the S-treatment have therefore to specify not only the agents' optimal effort choices but also their optimal sabotage decisions. It turns out that for $\Delta < 60$ both agents do not sabotage and choose efforts as in the B-treatment. For $\Delta \geq 60$ both agents should sabotage and should choose the lowest possible effort (see Table 2).

To derive this prediction first note that whenever agent i decides to sabotage agent j's output, it is optimal for agent i to choose the lowest possible effort level, i.e., $e = 1$. This simply holds because in case i sabotages agent j's output, y_j is zero while y_i is strictly larger than zero for all e_i. To construct an equilibrium in the S-treatment we start by assuming that nobody sabotages ($s_i = s_j = 0$). In this situation incentives are as in the B-treatment and efforts should be chosen accordingly. Let us denote the optimal effort level for $s_i = s_j = 0$ with $\hat{e}$. Given that both agents choose $\hat{e}$ and $s = 0$, we need to check whether agent i has a unilateral incentive to deviate, i.e., whether $EU_i(e_i = 1, s_i = 1; \hat{e}_j, s_j = 0) \geq EU_i(\hat{e}_i, s_i = 0; \hat{e}_j, s_j = 0)$. It can easily be calculated that there is no incentive to deviate for all $\Delta < 60$. As an example consider the case where $\Delta = 20$ and remember that sabotage costs are $c(s) = 27$. If both agents decide not to sabotage each other,

[6] It is necessary, however, to also check whether in equilibrium the expected utility is higher than w_l for all contracts. This is indeed the case for all treatments.

it is optimal to choose $\hat{e} = 2$. In this situation both earn an expected payoff of $EU_i = 0.5 \cdot 160 + 0.5 \cdot 140 - c(2) = 149$. Neither agent has an incentive to deviate and to sabotage, because if agent i sabotages, his payoff amounts to $160 - 27 = 133 < 149$.

For $\Delta \geq 60$, agent i has an incentive to deviate and to sabotage. The question now is whether agent j sticks to his strategy or whether he also has an incentive to deviate. Note first that whenever i sabotages the output of j, the latter has an incentive to choose $e_j = 1$ since his output is zero anyway. We therefore have to ask, whether $EU_i(e_i = 1, s_i = 1; \hat{e}_j = 1, s_j = 1) \geq EU_i(\hat{e}_i, s_i = 0; \hat{e}_j = 1, s_j = 0)$ in which case j would also sabotage. Simple calculation reveals that agent j does in fact have the incentive to sabotage i for $\Delta \geq 60$. Since i and j are symmetric this implies that for $\Delta \geq 60$ both agents should choose $e = 1$ and $s = 1$. To illustrate this prediction consider the case where $\Delta = 100$. Here, $EU_i = 125$ if both agents choose $\hat{e} = 6$ and do not sabotage each other. Now it is profitable to deviate for i and to choose $(ei = 1, si = 1)$, which yields a profit of $U_i = 200 - 27 = 173$. If j chooses $e_j = 1$ and $s_j = 0$ his payoff is 100. If he also sabotages, however, his expected payoff is $EU_j = 0.5 \cdot 200 + 0.5 \cdot 100 - 27 = 123 > 100$. Taken together for all contracts with $\Delta > 60$ there should be no sabotage and efforts should be chosen as in the B-treatment. For all $\Delta \geq 60$, both agents sabotage and choose $e = 1$. Given this relation between Δ and e, it is of course no longer optimal for principals to choose the maximum Δ. Since for $\Delta \geq 60$ both agents sabotage, expected profits are zero for $\Delta \geq 60$. Since up to $\Delta = 40$ effort is increasing in Δ it is optimal to choose $\Delta^* = 40$ in the S-treatment.

As outlined above, the theoretical prediction for the Baseline- and the Loss-treatment is exactly the same. The relationship between spreads and effort should be equal and, as a consequence, profit-maximizing principals should in both treatments choose the same (high) spreads. In the presence of loss averse agents, however, behavior might differ between the B- and the L-treatment. The fact that people are loss averse and try to avoid losses has been shown in various decision experiments, starting with the well known experiments presented by Kahneman and Tversky

(1979).[7] A simple way to characterize loss aversion is to assume that agents have a utility function that is linear in payoffs, except for a kink at the reference point such that the marginal utility of earnings drops discretely once the reference point is attained (Kahneman and Tversky (1979)).

To our knowledge, there is no evidence on whether varying the opportunities for losses affects behavior in a principal agent framework. To demonstrate how loss aversion can be relevant in our set-up, assume that both agents choose their equilibrium efforts e^*, and assume that a payoff of zero serves as a salient reference point for any given tournament interaction, so that agents attach a special importance to achieving at least zero in any given tournament interaction. The dislike for falling below zero reflects the fact that marginal utility of income is higher below the reference point than above. In the B-treatment (where the wage sum is 300) both agents, the winner as well as the loser of the tournament make a gain (positive profit), regardless of the level of the spread. In the L-treatment, however, if agents choose effort, e^*, according to the theory, the loser's payoff falls below the reference point of zero for sufficiently high spreads. Equilibrium profits for the winner and the loser, dependent on the spreads $(\pi_w(e^*), \pi_l(e^*); \Delta)$ are $(70, 70; 0)$, $(79, 59; 20)$, $(86, 46; 40)$, $(91, 31; 60)$, $(94, 14; 80)$, $(95, -5; 100)$, $(94, -26; 120)$ and $(91, -49; 140)$. Thus, for spreads above 80, the loser makes a negative payoff in equilibrium if she increases effort with the spread as predicted by theory. A possible strategy for a loss-averse subject in this setting is to choose the profit maximizing effort for each tournament interaction conditional on avoiding losses. Such an effort is equal to e^* for all spreads up to 80, but is lower than e^* for all spreads above 80.

Subjects in the experiment were told in the instructions that in case they incurred losses in a given transaction, these losses would be paid out of the accumulated income or if necessary the show-up fee. Even if a subject made a loss in every tournament interaction, these losses would be more than covered by the show up fee and thus negative earnings for the experiment were not possible.

[7] See also, e.g., Gneezy et al. (2003) for evidence on loss aversion in double auctions, and Andreoni (1995) for evidence from public good games.

A key question in the literature on reference-dependent preferences is how the reference point is determined. It seems likely that agents in our setting might consider each tournament interaction as a separate decision, and have a reference level in mind for a given tournament interaction. This would be consistent with evidence that people tend to bracket risky choices narrowly, rather than integrating individual choices into a larger compound lottery (Read et al., 1999; Rabin and Weizsaeker, 2007), and with evidence that people are myopically loss averse, in the sense of being loss averse around a reference point for individual lottery choices even when these are part of a sequence of such choices (e.g., Gneezy et al., 2003; Gaechter et al., 2007). We hypothesized that a particularly salient reference point for a tournament interaction is a payoff of zero, and thus designed the L-treatment to shift the loser prize below zero for larger spreads.

Ultimately, it is an empirical question addressed by our analysis, whether agent behavior in tournament interactions is best described by loss aversion around a reference point of zero. If this reference point is sufficiently salient, and aversion to losses sufficiently strong (i.e., the kink in utility at zero is sufficiently pronounced), an agent would be observed to deviate from the equilibrium effort and choose some $e < e^*$ precisely for those tournament interactions that involve spreads above 80, whereas in the absence of loss aversion, or if another value besides zero is the most salient reference point, this would not be the case.[8]

In summary, if sufficiently many subjects are loss averse around zero, we should in the L-treatment observe an effort-spread slope which is the same as in the B2-treatment for all $\Delta \leq 80$. For $\Delta > 80$, however, the effort-spread slope should be flatter in the L- than in the B2-treatment. For the wage setting policy of principals

[8] If utility is smoothly concave around zero, rather than kinked, with concavity increasing as zero is approached from the right, this could also generate loss aversion. Note that smoothly concave utility around zero would not be the product of standard risk aversion by a fully rational agent. As stakes are small relative to lifetime wealth, fully integrating outcomes in the experiment with wealth would lead to utility that is very close to linear (Rabin, 2000). Instead it seems to requires the narrow bracketing that is characteristic of the psychological evidence on reference-dependent decision making to explain why utility should be defined over outcomes in a given tournament interaction rather than final wealth levels. Regardless, our interest in this paper is not in determining whether utility is kinked or smoothly concave around zero, but rather in the implications of loss averse *behavior* for the design of tournament incentives.

this implies that for high spreads profits increase at a lower rate in the L- compared to the B2-treatment. It is even possible that profits do not increase or even decrease for spreads of 100, 120 or 140. If this holds, firms should therefore choose the maximum spread less often than in the B-treatment. As a consequence, one would expect wage compression in the L-treatment.

4 Results

4.1 The baseline treatment

We set the stage by analyzing behavior of principals and agents in our baseline treatment, where sabotage and opportunity for losses are ruled out by design. The findings, which are strongly supportive of the central predictions of tournament theory outlined in Section 3, can be summarized in the following result:

Result 1: *In a sparse environment principals choose high wage spreads. The tournaments create powerful incentives: effort is on average monotonously increasing in spread, almost exactly as predicted. Profits are highest for high spreads, making principals' choices quite rational.*

Support for R1 is provided by Figures 1 to 3. Figure 1 shows that in the baseline treatment principals chose predominantly the contracts with the highest wage spreads. The modal choice is the contract with the highest possible spread of 140, and of all contracts, 77 percent have a spread higher than 80. Notably, the choices appear very similar in the two treatments, suggesting that order effects or learning had little role to play. In fact, the non-parametric Mann Whitney U-test does not reject the hypothesis that the mean spread levels in the five B1 and the five B2 sessions are exactly the same ($p < 0.209$). The mean spreads are 108.3 in the B1-treatment and 113.2 in the B2-treatment.

[Figure 1 about here]

Figure 2 shows the mean effort choices in the B-treatment dependent on the spreads chosen by the principals. We plot both the B1- and the B2-treatments and also the theoretical prediction. The figure reveals two important findings. First of all, the agents' average effort decisions increase in the spread level and are surprisingly close to the theoretical prediction. For spreads between zero and 40, efforts are somewhat higher than predicted. Note however, that these spreads are very rare (see Figure 1). For spreads above 40, mean efforts are remarkably close to the prediction. Figure 2 thus reveals that agents respond very sensitively to the incentives they are exposed to. A second important insight from the figure is that agents' choices in B1 and B2 are almost identical, and are not significantly different when comparing session means for effort in the two treatments (Mann Whitney U-test, $p = 0.347$). Thus behavior is quite robust and does not significantly change with experience.

[Figure 2 about here]

In Figure 3 we see the average profits for principals, associated with different spread levels. It is evident that in the B-treatments, profits are the higher the stronger the incentives. Mean profits are about twice as high for spreads of 120 or 140 compared to spreads of zero or 20. In the light of the results shown in Figure 3, the wage policy of principals as shown in Figure 1 was quite rational. Profits are highest for spreads of 140, and the most frequently chosen spread in both B1 and B2 is in fact 140. The spearman rank correlation between mean profit for a given Δ and the frequency of that Δ is positive and highly significant in both B1 ($p < .0001$, two sided, n = 8) and in B2 (($p < .0002$, two sided, n = 8).

[Figure 3 about here]

4.2 Impact of sabotage

Our second question concerns the impact of sabotage on the effectiveness of tournament incentives, and the corresponding contract choices of prinicpals. We first investigate whether the possibility to engage in sabotage activity induces wage compression compared to the B1-treatment, where sabotage was ruled out by design. The evidence on this question is summarized in

Result 2: *The possibility to engage in sabotage causes wage compression. Principals choose much more egalitarian wages in the S- than in the B1-treatment.*

Support for Result 2 comes from Figure 4 which shows the relative frequency of chosen wage spreads in the B1- and the S-treatments. The figure reveals that the spread distributions are very different. In the B1-treatment, about 75 percent of the contracts involve spreads of 100 or above. Quite to the contrary, almost 90 percent of the wage spreads in the S-treatment are 60 or below. The average wage spread in the B1-treatment is 108.3 while it is only 29.4 in the S-treatment. Comparing the five session means of the B1- and the fives session means of the S-treatment confirms that the difference between spreads is statistically significant ($p = .043$, Wilcoxon signed rank test, two-sided). Thus, the same principals who choose high incentives in the B-treatment rely on rather weak incentives in the S-treatment.

[Figure 4 about here]

Result 2 lends strong support to the theoretical predictions derived in Section 3. Principals do in fact choose relatively high spreads in the B- and relatively low spreads in the S-treatment. Theoretically this difference is driven by the difference in effort behavior. While in the B-treatment efforts and consequently profits should increase in the spread level, high spreads in the S-treatment are predicted to lead to sabotage and low profits. Our next result supports this prediction:

Result 3: *In contrast to the B-treatments, in the S-treatment efforts first increase in the spread level but then drop to low levels. Sabotage activity is low for low spreads and high for spreads above 20.*

Support for R3 comes from Figures 5 to 6. Figure 5 shows the mean effort levels in the S-treatment dependent on the spreads, together with the theoretical prediction.

A first observation is that mean effort levels are much lower compared to in the B-treatment as predicted (see Figure 2). Second, the figure shows that the theoretical prediction regarding effort levels and wage spreads is qualitatively confirmed: Efforts are about 1 for a spread of zero in the S-treatment, then increase to about 2 for spreads of 20, and then are slightly above 1 for spreads larger than 40. It is interesting to note that mean efforts are highest at a spread of 20 and not at a spread of 40. For $\Delta = 40$, mean efforts are not only lower than predicted but also lower than in the B1-treatment. A possible explanation for this finding is the sabotage behavior to which we turn now.

[Figure 5 about here]

Sabotage occurred in 194 of the possible 888 cases (22 percent). Figure 6 shows that the sabotage frequency is very low for spreads of zero and 20 but then strongly increases in the spread level.[9] The average sabotage rate for spreads between 80 and 140 is 88 percent. Thus sabotage behavior is in line with the theoretical prediction, except that for spreads of 40 the sabotage rate is not zero but 40 percent. This excess sabotage also helps to explain why mean efforts are rather low for spreads of 40. It is not rational to sabotage and at the same time to put forward an effort higher than 1. In fact in 91.2 percent of the cases where agents sabotage, they choose e = 1 with a resulting mean effort of e = 1.13. If we exclude the subjects who sabotage, mean effort at a spread of 40 is higher than at a spread of 20. But even if we consider only the effort choices of the subjects who do not sabotage, mean efforts are lower in the S- than in the B1-treatment. This holds for all spread levels. Therefore, we conclude that even those subjects who do not engage in sabotage activities are affected by the (expected) sabotage of others. Thus, the possibility of sabotage has two effects on effort behavior: a direct effect, which simply means that those who sabotage choose $e = 1$, and an indirect effect which implies lower effort choices by those who do not sabotage but who expect to be the possible victim of sabotage by others.

[9] The relative frequency of the spread 120 is extremely low (see Figure 4).

[Figure 6 about here]

The effort and sabotage behavior of agents determines the profitability of different wage policies by principals. Given this profitability we can address the question, whether the wage policies shown in Figure 4 were optimal. The main findings are summarized in

Result 4: *In contrast to the B1 treatment, in the S-treatment profits first increase in the wage spread but then drop to very low levels. Given this profit-spread relation, principals' wage policies were quite rational.*

Empirical evidence for R4 is given in Figure 7, which shows the mean profits of principals in the B1- and the S-treatment. Whereas in the B1-treatment profits are the higher the stronger the incentives, in the S-treatment wage compression pays. Mean profits in the S-treatment are highest for modest spreads with a maximum at a spread of 20. Since higher spreads provoke extensive sabotage it does not pay for firms to set stronger incentives. In the light of the results shown in Figure 7, the wage policy of principals as shown in Figure 4 was quite rational. In the S-treatment the most profitable spread is 20, and this is by far the most frequently chosen spread. The spearman rank correlation between mean profit of a spread and its frequency is also positive and highly significant ($p < 0.05$, two sided, n = 8).

[Figure 7 about here]

4.3 Loss aversion and tournaments

We now turn to a comparison between the L-treatment and the B2-treatment. As we have discussed in Section 3, the theoretical prediction in the absence of loss aversion effects is exactly the same for both treatments: The relationship between spreads and efforts should be equal and, as a consequence, profit-maximizing principals should in both treatments choose the same (high) spreads. If sufficiently many subjects are loss averse, however, we should in the L-treatment observe an effort-spread slope, which is indistinguishable from that in the B2-treatment for $\Delta \leq 80$, while for $\Delta > 80$, the effort-spread slope should be flatter in the L- than in the B2-treatment. This might imply that wages are more compressed in the L- compared to the B2-treatment. By comparing the results from the L- and the B2-treatment we can directly test both predictions. Our main finding concerning the principals wage policy is summarized in

Result 5: *There is wage compression in the L-treatment.*

Result 5 is supported by Figure 8, which shows the distribution of wage spreads in the L- and the B2-treatment. The figure clearly indicates that principals rely on strong incentives in the B2-treatment where the most frequently chosen spread is 140. To the contrary, principals prefer intermediate spreads in the L-treatment. The mean wage spread is 113.2 in the B2-treatment and 84.7 in the L-treatment. Comparing the five session means of the L- and the B2-treatment confirms that the differences in the spreads are statistically significant (p = .043, Wilcoxon signed rank test, two-sided). Thus the same principals who choose relatively high spreads in the B2-treatment select significantly lower spreads in the L-treatment.

[Figure 8 about here]

Is the different wage setting policy shown in Figure 8 the response to a different effort behavior of agents across treatments? Our next result provides evidence in favor of this interpretation.

Result 6: *The spread-effort relation is weaker in the L- than in the B2-treatment. This holds in particular for high spreads, suggesting that agents are loss averse.*

Result 6 will be discussed primarily with the help of Table 3. In this table we show three OLS-regression models that investigate the spread-effort relationship in the L- and the B2-treatment. Model 1, shown in Columns (1) and (2), simply regresses effort on spreads controlling for possible period effects. In both treatments the spread coefficient is positive and highly significant. It is larger, however, in the B2-treatment. This means that the same agents supply lower efforts for a given spread in the L- than in the B2-treatment.[10] Remember that we have hypothesized that if agents are loss averse, the effect of loss aversion should be visible in particular for higher spreads. Up to a spread level of 80, agents can choose e^* without facing a positive loss probability. We therefore expected that up to a level of 80, effort responses between the L- and the B2-treatment should not significantly differ, even if agents are loss averse. For higher spreads, however, efforts should be higher in the B2- than in the L-treatment. To test this hypothesis we set up regression models 2 and 3.

[Table 3 about here]

In Model 2 effort is regressed on low spreads (including all spreads smaller or equal 80) and on high spreads (including all spreads above 80).[11] The results are quite remarkable. The coefficient for lower spreads is roughly 0.05, which is exactly the slope predicted by standard economic theory. It is also highly significant. For higher spreads, the coefficient is not only much smaller (by a factor of five) but also highly insignificant. This suggests that as long as losses are impossible, agents behave as predicted by standard economic theory. If losses become possible, however, loss averse subjects react by choosing lower efforts than predicted. This interpretation is supported by the results for the B2-treatment. Here the two spread coefficients are significantly positive and almost identical. Agents in the B2-treatment react smoothly to an increase of incentives, regardless of the level of the spread. It is also interesting to see that for low spreads the coefficients in the L- and the B2-treatment are very similar while they are very different for high spreads.

To test whether the coefficients for low and high spreads are in fact different, we set up Model 3. In this regression model we simultaneously estimate effort-spread

[10] Note also that mean effort is higher in the B2- than in the L-treatment (5.58 in the L- and 6.84 in the B2-treatmnt). A comparison between the five session means in the L- and the B2-treatment reveals that the difference is significant ($p < 0.05$, Wilcoxon signed rank test, two sided).

[11] The constant, therefore, measures the mean effort at a spread level of 80.

coefficients for both treatments. The coefficients *Low Spreads* and *High Spreads* measure the slope in the L-treatment. In the rows below we introduce two interaction terms, where the spreads are interacted with a treatment dummy for B2. These two coefficients therefore measure the difference in the slopes between the L- and the B2-treament. Model 3 supports the previous findings: In particular, the coefficient of the interaction term *Low Spreads x B2* is insignificant, which implies that for low spreads effort responses are basically the same in the L- and the B2-treatment. For spreads above 80, efforts are significantly higher in the B2- than in the L-treatment as revealed by the positive and significant coefficient of the interaction term *High Spreads x B2*. The negative coefficient *B2-treatment* indicates that at $\Delta = 80$, efforts are somewhat lower in the B2- than the L-treatment.

Support for the relevance of loss aversion comes also from an analysis of individual data. Remember that there is a total of 74 agents who participated in the L- and the B2-treatment. All agents were exposed to different spreads. Thus we can calculate for each agent two correlation coefficients between spread and effort, one for the L- and one for the B2-treatment. If there is loss aversion, the coefficients should be smaller in the L- than in the B2-treatment, at least for some of the agents. Comparing the 148 coefficients reveals that this is in fact the case (Wilcoxon signed-rank test, $p = 0.0619$, two sided).

The effort patterns in the B2- and the L-treatment help to understand the wage policy of principals: In the B2-treatment, efforts are increasing in the spread level for all spreads, and as a consequence so are profits for principals (see Figure 3. Therefore, principals' policy of choosing high spreads (see Figure 8) is understandable. In the L-treatment, we have shown strong evidence for loss aversion effects both on an aggregate and on an individual level, with implications for principals' profits. As the regression results of Table 3 indicate, average efforts are increasing up to a spread level of 80 and are only insignificantly rising for spreads above 80. Figure 9 shows that profits for principals are increasing monotonically in the spread until an abrupt change once spreads exceed 80. Profits are lower for spreads of 100, and about the same for 120, compared to spreads of 80. Profits are slightly higher for spreads of 140 than for 180. This contrasts with B2-treatment, for which profits increase in the spread for all spreads, and for which there is a steep increase in profits as spreads increase above 80. Moreover, an analysis of the variance of efforts reveals that for spreads above 80, the variance as well as the coefficient of variation actually increases in the spread level in the L-treatment.[12] If principals are even slightly risk averse they might strictly prefer a spread of 80 over a spread of 100, 120 or 140. This way they face a lower variance without loosing anything in terms of expected efforts. Even if there is not a strict preference for spreads of 80, the weaker link between high spreads and profits helps explain why principals are less likely to choose the highest spreads than in baseline (Figure 8). It follows that the more compressed spread distribution in the L-treatment compared to the B2-treatment makes sense. The optimality of the chosen contracts is further supported by the fact that in the L-treatment, the spearman rank correlation between mean profits and frequencies of spreads is again positive and significant ($p = 0.0580$, two sided, n = 8), showing that principals choose significantly more often those spreads that are more profitable.

5 conclusion

This paper approaches the study of tournament incentives from a different angle than most of the empirical literature, investigating how principals choose to structure tournaments depending on the institutional setting. We find that the way that principals design tournament incentives depends on details of the work environment, as well as on their understanding of the psychological factors or preferences that play a role in the way that agents respond to incentives.

One implication of our results is that principals will choose large or small prize spreads, depending presence of sabotage opportunities in the workplace. Consistent

[12] The increase in variance might have to do with various types of agents, those who are loss averse and those who are not. In fact, the comparison with the B2-treatment shows that in B2-treatment the coefficient of variation is smaller than in the L-treatment for spreads of 100, 120 and 140.

with the predictions of Lazear (1989), principals favor greater wage equality in work settings with opportunities for sabotage, because increasing the prize spread causes agents to shift from productive activity to sabotage of opponents. Sabotage is even slightly more pronounced than predicted based on the material incentives in the experiment. Although speculative, this could potentially reflect pre-emptive sabotage, due to a non-material motive such as betrayal aversion. Recent findings suggest that people are particularly averse to betrayal, or being a sucker (Bohnet and Zeckhauser, 2004). This type of motive could potentially lead people to sabotage too early, to prevent the possibility that someone else sabotages them and they end up being the sucker. Betrayal aversion could be a factor that exacerbates the problems of sabotage in the workplace.

Another key finding is that a principal's choice of tournament incentives may depend on the prevalence of loss averse agents in the workforce, and on whether a large prize spread can be employed without causing losers to fall below a reference payoff. Constraints on the wage sum available to a principal are one factor that can make it so that large prize spreads do entail losses for agents, because raising the spread requires driving the loser prize down below the reference point. Agents respond to increasing prize spreads as predicted by theory, until the increasing spread entails a prize below the reference point, at which point further increases in the spread are ineffective. Thus, in the presence of loss averse agents, the wage sum may have an impact on the design of tournament incentives, by making losses possible for large prize spreads and causing principals to compress wages accordingly. An important implication is that the success of tournament incentives might be enhanced, if principals are able to manage agents' reference points, and thus affect the way that agents view the outcome of losing the tournament.

Although agent effort levels increase monotonically in the prize spread in our baseline treatment, as predicted by theory, effort provision tends to be a bit higher than predicted for almost every prize spread (see Figure 2). One possible explanation, which could be explored in future research, is envy, or aversion to disadvantageous inequity in payoffs (see Fehr and Schmidt, 1999). Our design minimizes the scope for payoff comparisons between agents and the principal by making the principal's payoff function unknown, a feature that is probably realistic in many settings where agents do not have a firm idea of the principal's payoff, or where agents view the principal as not be a relevant comparison figure. There are almost necessarily possibilities for lateral comparisons between agents in a tournament setting, however, given that tournament incentives generally involve agents having at least approximate knowledge of the relevant prizes. Such comparisons are possible in all our our experiment treatments. As shown by Grund and Sliwka (2005), inequity aversion can increase an agent's effort in a tournament setting, because of the desire to avoid having a relatively lower payoff compared to the other agent (this holds under the reasonable assumption that disadvantageous inequity aversion is stronger than advantageous inequity aversion). Their model assumes inequity judgements are based only on prize payments, and do not take into account effort costs. If agents instead evaluate inequity based on payments net of effort costs, an agent's decision is made more complicated by the fact that although working hard increases the possibility of winning the tournament, it also worsens the relative payoff in the case that the agent loses. But if the former effect dominates, then disadvantageous inequity aversion could generate systematically higher efforts than predicted by the standard model. This type of fairness comparison is likely to be relevant in most types of tournaments, given that a policy of wage secrecy would defeat the purpose of the tournament. Furthermore, to the extent that inequity aversion increases agents' efforts, principals would want to actively encourage lateral comparisons in tournament settings.

References

Andreoni, J. (1995): "Warm-glow versus cold-prickle: The effects of positive and negative framing on cooperation in experiments," *Quarterly Journal of Economics*, 110(10), 1–21.

Becker, B., and M. Huselid (1992): "The incentive effects of tournament compensation systems," *Administrative Science Quarterly*, 37, 336–350.

Bull, C., A. Schotter, and K. Weigelt (1987): "Tournaments and piece rates: An experimental study," *Journal of Political Economy*, 95(1), 1–33.

Ehrenberg, R., and M. Bognanno (1990): "The incentive effects of tournaments revisited: Evidence from the European PGA tour," *Industrial Labor Relations Review*, 43, 74–89.

Eriksson, T. (1999): "Executive Compensation and Tournament Theory: Empirical Tests on Danish Data," *Journal of Labor Economics*, 17, 262–280.

Fehr, E., and A. Falk (1999): "Wage rigidity in a competitive incomplete contract market," *Journal of Political Economy*, 107, 106–134.

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

Fischbacher, U. (2007): "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics*, 10(2), 171–178.

Gächter, S., A. Herrmann, and E. Johnson (2007): "Individual-Level Loss Aversion in Riskless and Risky Choices," CeDEx Discussion Paper No. 2006-02.

Gibbons, R. (1992): *Game Theory for Applied Eonomists*. Princeton University Press, Princeton.

Gneezy, U., A. Kapteyn, and J. Potters (2003): "Evaluation Periods and Asset Prices in a Market Experiment," *Journal of Finance*, LVIII(2), 821–837.

Grund, C., and D. Sliwka (2005): "Envy and Compassion in Tournaments," *Journal of Economics and Management Strategy*, 14, 187–207.

Harbring, C., and B. Irlenbusch (2005): "Incentives in Tournaments with Endogenous Prize Selection," *Journal of Institutional and Theoretical Economics*, 161, 636–663.

——— (2008): "How many winners are good to have? On tournaments with sabotage," *Journal of Economic Behavior and Organization, forthcoming.*

Kahneman, D., and A. Tversky (1979): "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 407, 263–291.

Knoeber, C., and W. Thurman (1994): "Testing the theory of tournaments: An empirical analysis of boiler production," *Journal of Labor Economics*, 12(2), 155–179.

Lazear, E. (1989): "Pay equality and industrial politics," *Journal of Political Economy*, 97(3), 561–580.

Lazear, E., and S. Rosen (1981): "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89(5), 841–864.

Nalebuff, B., and J. Stiglitz (1983): "Prizes and Incentives: Towards a General Theory of Compensation and Competition," *The Bell Journal of Economics*, 14(1), 21–43.

Rabin, M. (2000): "Risk Aversion and Expected Utility Theory: A Calibration Theorem," *Econometrica*, 68(5), 1281–1292.

Rabin, M., and G. Weizsaecker (2007): "Narrow Bracketing and Dominated Choices," Working paper, London School of Economics.

Read, D., G. Loewenstein, and M. Rabin (1999): "Choice Bracketing," *Journal of Risk and Uncertainty*, 19(1-3), 171–197.

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

Table 1: Different contract types

	Wage for worker with higher output w_h	Wage for worker with lower output w_l	$\Delta = w_h - w_l$
1	150	150	0
2	160	140	20
3	170	130	40
4	180	120	60
5	190	110	80
6	200	100	100
7	210	90	120
8	220	80	140

Table 2: Summary of predicted effort and sabotage levels

$\Delta = w_h - w_l$	B-treatment	S-treatment	L-treatment
0	$e^* = 1$	$e^* = 1, s^* = 0$	$e^* = 1$
20	$e^* = 2$	$e^* = 2, s^* = 0$	$e^* = 2$
40	$e^* = 3$	$e^* = 3, s^* = 0$	$e^* = 3$
60	$e^* = 4$	$e^* = 4, s^* = 1$	$e^* = 4$
80	$e^* = 5$	$e^* = 5, s^* = 1$	$e^* = 5$
100	$e^* = 6$	$e^* = 6, s^* = 1$	$e^* = 6$
120	$e^* = 7$	$e^* = 7, s^* = 1$	$e^* = 7$
140	$e^* = 8$	$e^* = 8, s^* = 1$	$e^* = 8$
Firm's payoff Maximizing Contract	$\Delta^* = 140$	$\Delta^* = 40$	$\Delta^* = 140$

Table 3: Effort responses in L- and B2-treatments

	Dependent variable: Effort				
	Model 1		Model 2		Model 3
	L-treatment	B2-treatment	L-treatment	B-treatment	L- and B2-treatments
Delta	0.030***	0.042***			
	[0.006]	[0.005]			
Period	-0.052	-0.055	-0.064	-0.055	-0.059***
	[0.041]	[0.028]	[0.035]	[0.028]	[0.011]
Low spread			0.052***	0.043***	0.051***
			[0.007]	[0.007]	[0.007]
High spread			0.012	0.042***	0.012
			[0.010]	[0.007]	[0.010]
Low spreads x B2					-0.009
					[0.010]
High spreads x B2					0.031**
					[0.011]
B2-treatment					-0.645**
					[0.228]
Constant	3.391***	2.416**	6.512***	5.805***	6.479***
	[0.234]	[0.638]	[0.446]	[0.303]	[0.301]
Observations	888	888	888	888	1776
Prob. > F	0.0024	0.0004	0.0025	0.0001	0.0675
R-squared	0.15	0.19	0.17	0.19	0.21

Notes: *Low spreads* are all spreads ≤ 80, *High spreads* are all spreads > 80. The estimation procedure is an OLS-regression with robust standard errors (in brackets) clustered on sessions (number of clusters per treatment $= 5$). ***, **,* indicate significance at 1-, 5-, and 10-percent level, respectively.

Figure 1: Contract choice in baseline

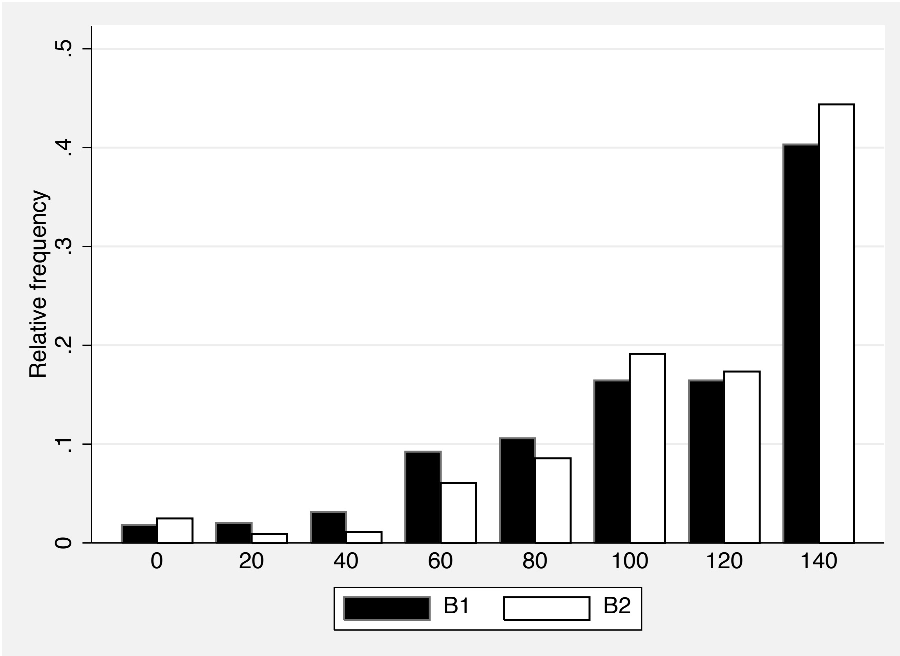


Figure 2: Effort choices in baseline

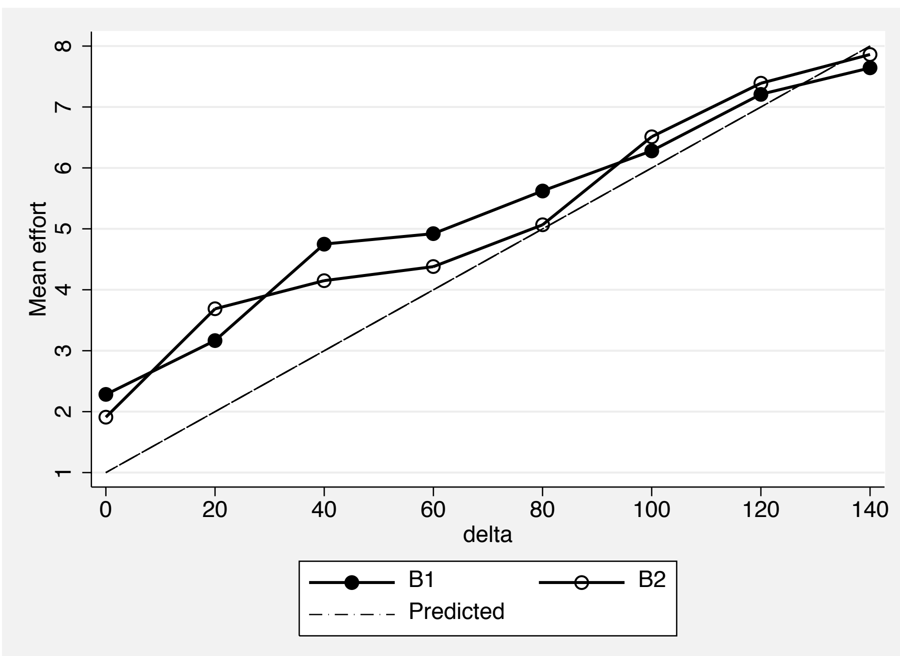


Figure 3: Profits for principals in baseline

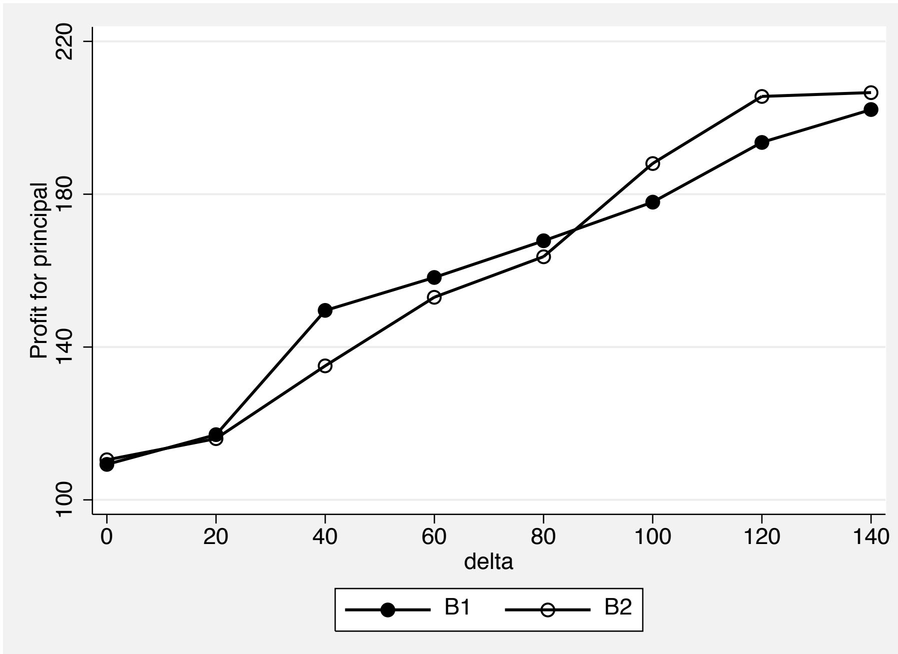


Figure 4: Sabotage and contract choice

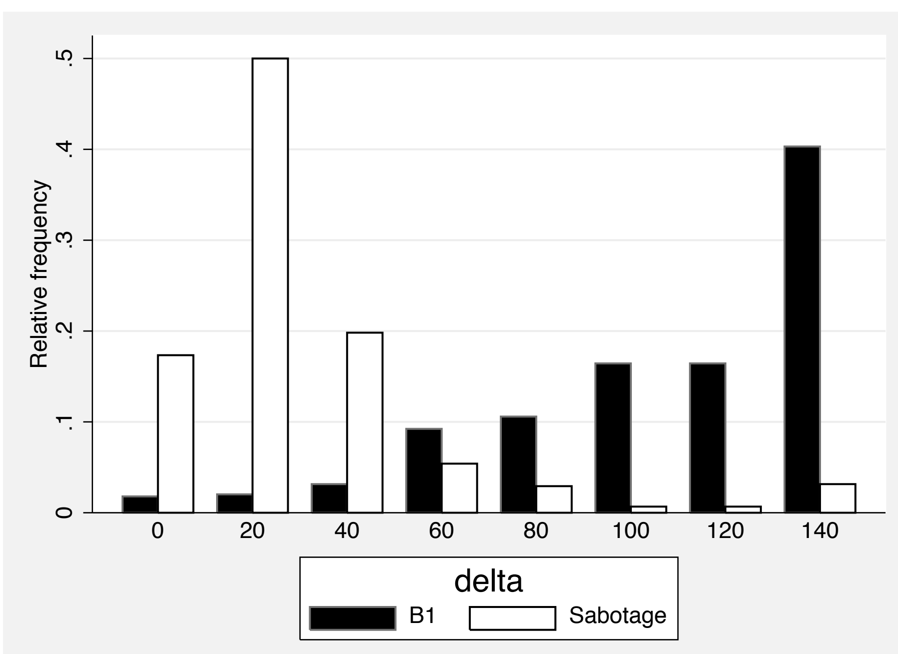


Figure 5: Effort in an environment with sabotage

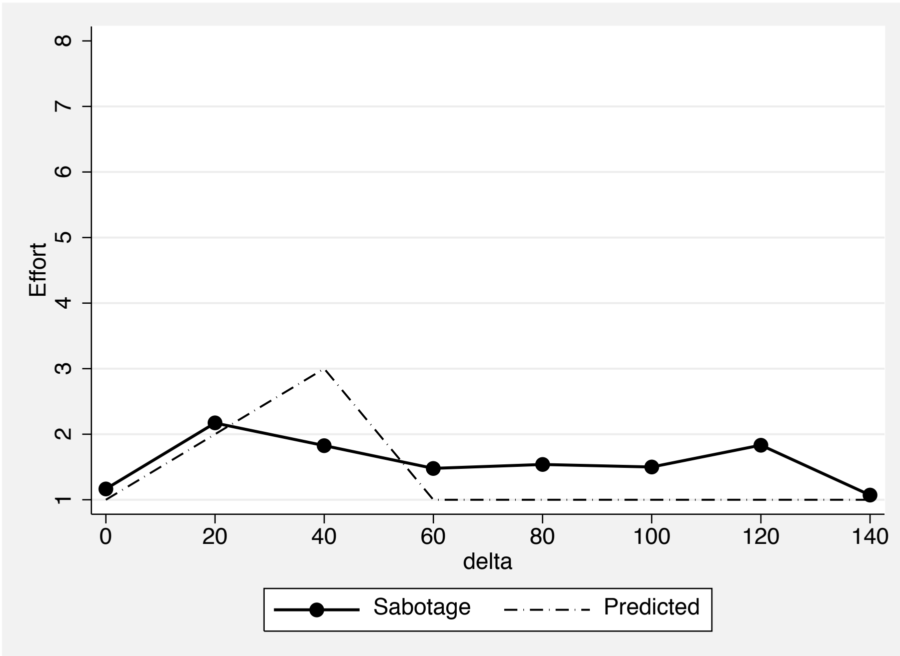


Figure 6: Sabotage frequency

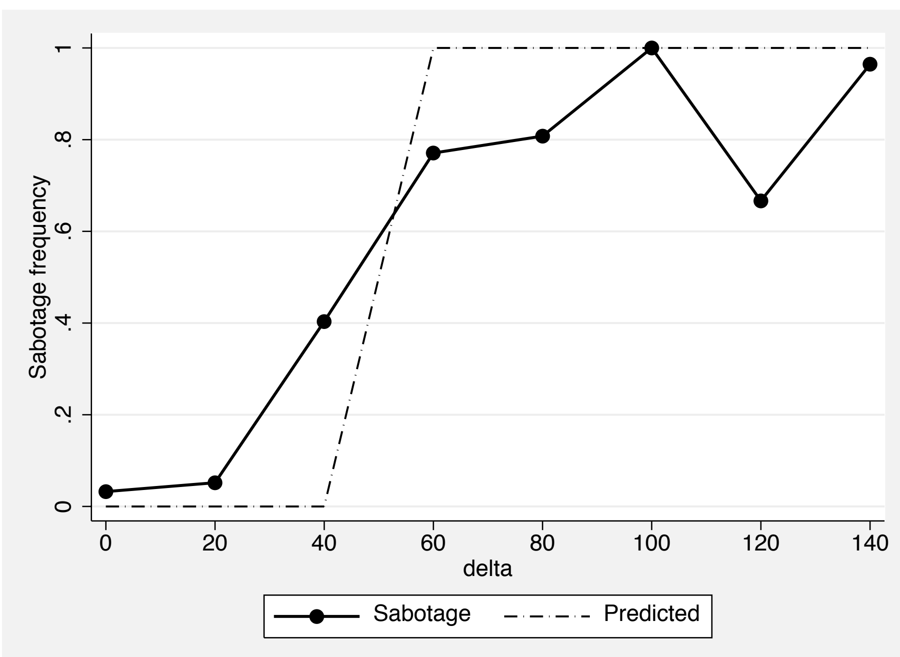


Figure 7: Sabotage and profits of principals

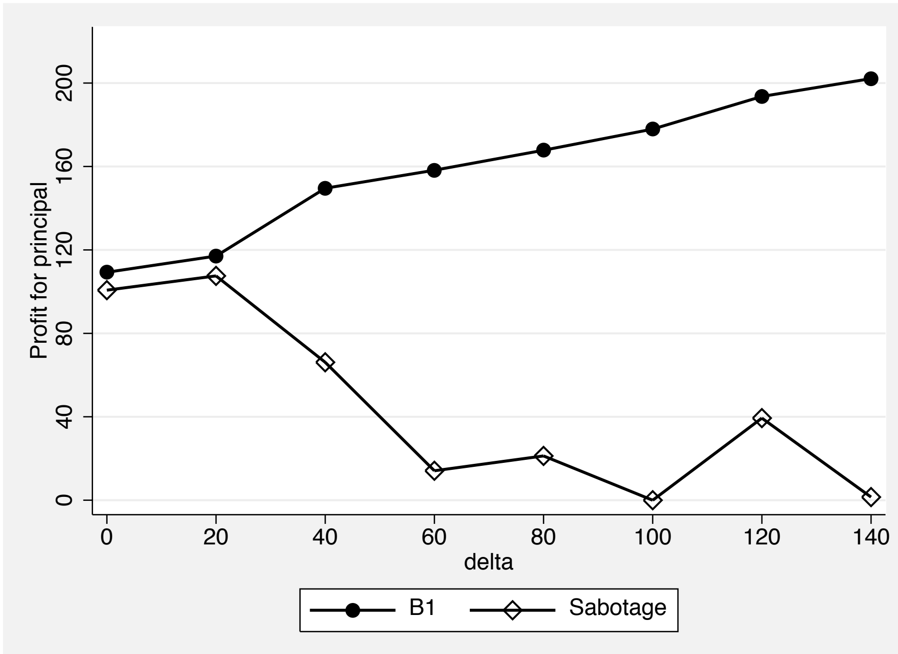


Figure 8: Loss aversion and contract choice

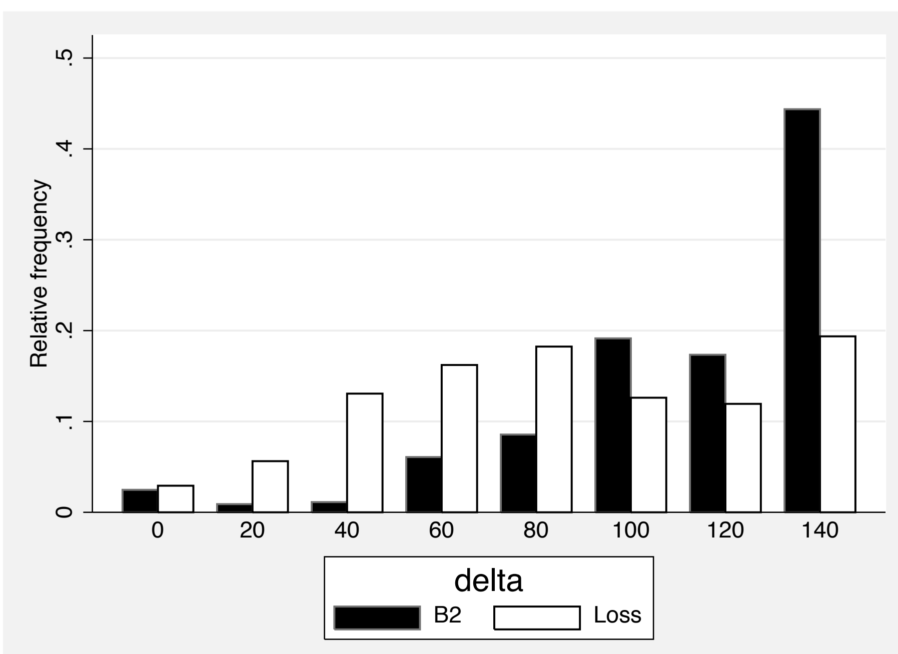
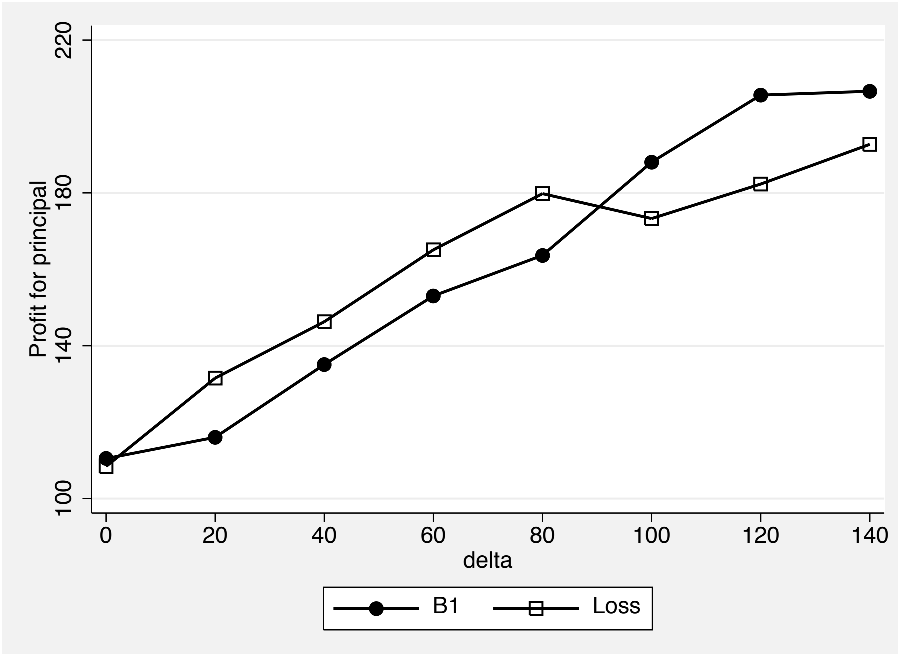


Figure 9: Loss aversion and profits of principals



APPENDIX: INSTRUCTIONS

In the following we present the instructions for the Baseline-treatment (for employers and employees). The Sabotage-treatment instructions are exactly the same with the only difference that employees had an additional choice, namely to "reduce the other employee's output to zero, at a cost of 27 Guilders". The instructions for the Loss-treatment are identical to those of the Baseline-treatment, except that the wage sum of the eight contract types is not 300 Guilders but 140 Guilders.

Instructions for Employers

You are now taking part in an economic experiment, which is financed by various research foundations. Please read the following instructions very carefully. If you have questions, please don't hesitate to ask us. Your question will be answered at your seat.

The instructions are intended solely for your personal information. **During the experiment, communication is not allowed.** Any violation of this rule will lead to the exclusion both from the experiment and from all payments.

During the experiment we talk about points instead of Francs. Your earnings will, therefore, be calculated in Guilders. At the end of the experiment the total amount of points you earn will be converted into Francs at the rate of

100 Guilders = 0.12 Francs.

On completion of the experiment you will receive **in cash** the number of Guilders earned during the experiment plus a 10 Francs show-up fee.

The following pages explain the experiment in detail.

Brief Survey on the Experiment's Procedure

There are 8 employers and 16 employees. During the entire experiment you are an **employer**.

In each period of the experiment two employees are matched with one employer. The participants are, therefore, in each period in a group of three. After each period new groups of three are randomly formed. **In each period you are, therefore, in a group with new participants.** You will never be informed about the other group members' identity.

Each period consists of **two stages**:

Stage one

At the beginning of **stage one** the employer chooses a type of contract, which determines the wage level for the two employees. The employer determines both the wage of the employee with the higher **work output**, as well as the wage of the employee with the lower work output. The higher the employees' work outputs are, the higher is the employer's income.

Stage two

After the employer has made its decision on the type of contract, **at stage two** both employees decide on their **effort**. An employee's work output depends directly on the chosen effort. The work output depends, furthermore, on an element of luck. The employees have to bear costs when choosing effort

When both the employer and the employees have made their choice, the participants are informed about the resulting incomes. This ends a period.

There is a total of 12 periods.

The incomes from all periods are added up, converted into Francs and paid out in cash on completion of the experiment.

The Experiment's Procedure in detail

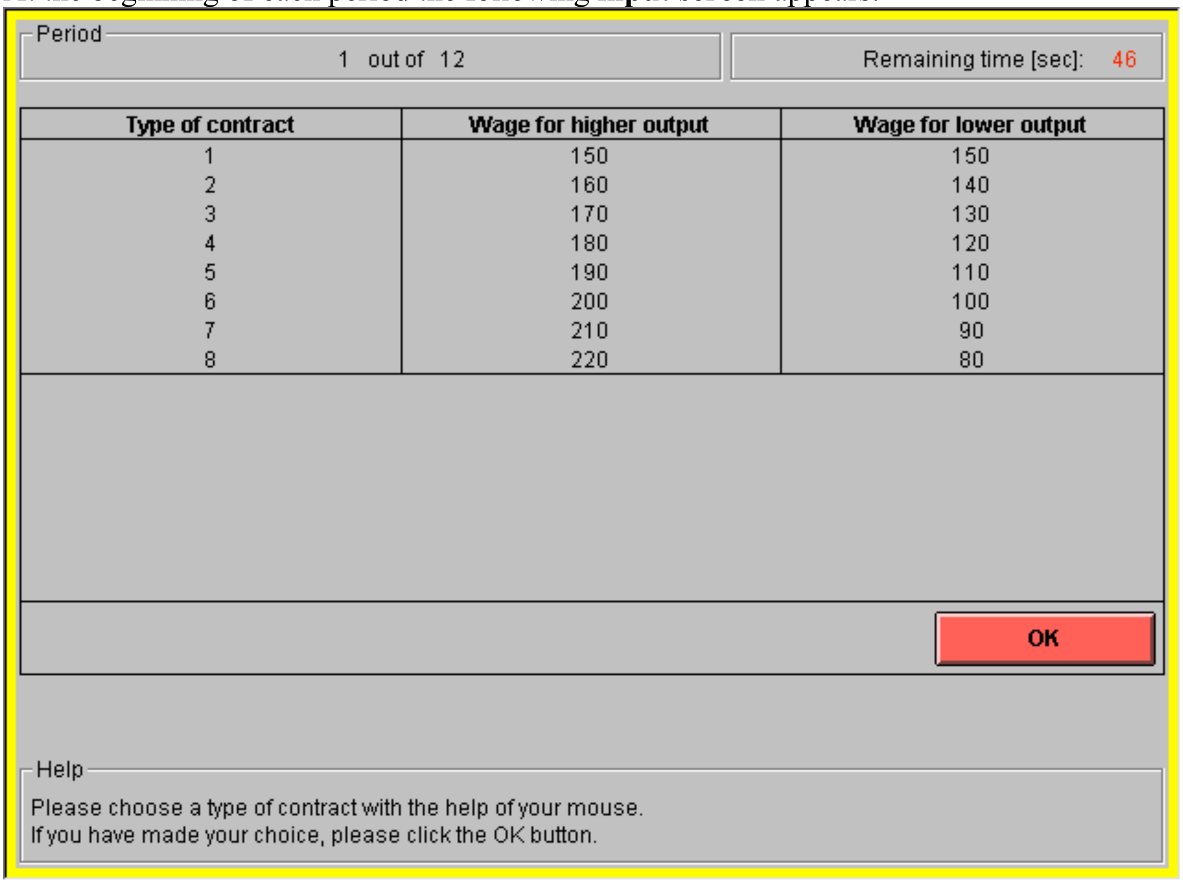
During the entire experiment you are an employer. In each of the 12 periods you are matched with two employees. The participants are divided into groups of three for the duration of each period. In each period new groups of three are randomly created. **Accordingly, in each of the twelve periods you are matched in a group with different participants.**

<u>Stage one</u>

At the beginning of each period you are **endowed with 300 Guilders**. Out of this endowment you pay the employees' wages. **The 300 Guilders are always completely spent for the payment of the wages.**

However, it is your decision **how** to divide these 300 Guilders between the two employees. This depends on which **contract types** you choose on the input screen. Your choice determines the employee's wage with the **lower work output**, as well as the wage of the employee with the **higher work output**.

At the beginning of each period the following **input screen** appears:



Period
1 out of 12
Remaining time [sec]: 46

Type of contract	Wage for higher output	Wage for lower output
1	150	150
2	160	140
3	170	130
4	180	120
5	190	110
6	200	100
7	210	90
8	220	80

OK

Help
Please choose a type of contract with the help of your mouse.
If you have made your choice, please click the OK button.

In the top left corner of the screen you see in which period you are. Below you see which types of contract you can choose. As you can infer from the table you can either pay both employees the same wage (contract type 1) or you pay them different wages (contract types 2 to 7). Please note that the wage sum is always the same. You choose a type of contract by clicking with the mouse at the corresponding row. After your choice you have to click the OK button. As long as you have not clicked the OK button, you can revise your decision.

As an example, if you decide on contract type 3 you pay the employee with the higher work output a wage of 160 Guilders and the employee with the lower work output a wage of 140 Guilders. A choice of contract type 1 pays both employees the identical wage of 150 Guilders. In this case both employees receive the same wage regardless of their work output.

Stage two

The employees' **work outputs** depend on the chosen **effort** as well as on a **random number**. The higher their effort and the higher the random number, the higher is their work output (see below).

At the beginning of stage two the employees decide on their effort, as soon as you have determined the type of contract. Both employees are aware of the type of contract you selected. They can choose an effort level between 1 and 12. Their effort choice is costly. The higher their effort is, the higher are the costs they will have to bear. The table below shows the feasible efforts and the corresponding costs. In general, the higher the effort, the higher are the costs involved.

Feasible efforts and effort costs:

effort	**1**	**1.5**	**2**	**2.5**	**3**	**3.5**	**4**	**4.5**	**5**	**5.5**	**6**	**6.5**	**7**
effort cost	0	0.25	1	2.25	4	6.25	9	12.25	16	20.25	25	30.25	36

effort	**7.5**	**8**	**8.5**	**9**	**9.5**	**10**	**10.5**	**11**	**11.5**	**12**
effort cost	42.25	49	56.25	64	72.25	81	90.25	100	110.25	121

Feasible are all efforts between 1 and 12 in steps of 0,5.

In addition to the effort, an employee's work output depends on a **random number**. For each employee, the computer generates a random number between 0 and 10. This random number is rounded up to four decimal places (i.e.: 0.4412; 8.1797; 2.0000; ....). Each number (between 0 and 10) is drawn by the computer with the same probability.

As soon as the employees have determined their effort and the computer has generated their respective random numbers, the work outputs are determined:

An employee's **work output** is calculated as follows:

Employee's work output = effort + random number

Thus, an employee's work output is the higher, the higher the effort and the higher the random number.

As an example, if an employee decides on an effort of 6 and the computer draws a random number of 3.7900 for him, his work output will amount to 9.7900. As an employer you will only learn about the work output. In this example you will learn that the employee's work output is 9.7900.

How to calculate the employees' income?

An employee's income depends on **whether he gets the higher or the lower wage and on the cost of effort**. The employee with the higher output gets the higher wage and the other employee gets the lower wage. The income of an employee is simply his wage minus the effort costs.

Employee's income = wage − effort costs

Thus, the income an employee is the higher, the higher the wage and the lower his chosen effort.

How do you calculate your income?

Your income depends on the work outputs of both employees. Your income is determined by the following formula:

Your income = **8 * (work output of one employee + work output of the other employee)**

As a result, your income is the higher, the higher the work outputs of both employees. The higher the employees' efforts and the higher the random numbers drawn for them, the higher is your income.

The work outputs of **both employees are added up and multiplied by 8**. Thus, your income is 8 times the work output of one employee plus 8 times the work output of the other employee.

If, for instance, one of the employees achieves a work output of 7.2100 and the other employee one of 4.7901, your income is (8 * 7.2100 + 8 * 4.7901) = 96.0008 Guilders.

When the employees have made their choice, an **income screen** informs you about both employees' work outputs and about your income in the current period.

As soon as you click the "continue" button on the income screen, a period is concluded, and the next period starts.

There is a total of 12 periods.

Are there any questions?

Control Exercises:

Please answer all questions and write down the whole calculation! If you have any questions please ask us!

1. Employee 1 has a work output of 7.5000; employee 2 one of 6.7000.

   What is your (the employer's) income? ....

2. Employee 1 chooses an effort of 4, employee 2 one of 3.5. The random numbers generated by the computer are 8.0000 and 1.8000 respectively.

   What is the work output of employee 1? ....

   What is the work output of employee 2? ....

   What is your income? ....

3. You choose the following contract type: The wage for the higher work output is 160, the wage for the lower work output is 140.

   What wage results for employee 1 of problem No. 2? ....

4. Both employees choose the minimal effort. Employee 1 gets a random number of 3.0000 and employee 2 a random number of 5.0000.

   What is your income? ....

5. The employer decides to pay the employee with the higher work output a wage of 210, and the employee with the lower work output a wage of 90. Employee 1 chooses an effort of 7.5 and employee 2 an effort of 8. The computer's generated random numbers are 2.0000 and 3.0000, respectively.

   What is the income of employee 1? ....

Instructions for Employees

You are now taking part in an economic experiment, which is financed by various research foundations. Please read the following instructions very carefully. If you have questions, please don't hesitate to ask us. Your question will be answered at your seat.

The instructions are intended solely for your personal information. **During the experiment, communication is not allowed.** Any violation of this rule will lead to the exclusion both from the experiment and from all payments.

During the experiment we talk about points instead of Francs. Your earnings will, therefore, be calculated in Guilders. At the end of the experiment the total amount of points you earn will be converted into Francs at the rate of

100 Guilders = 0.12 Francs.

On completion of the experiment you will receive **in cash** the number of Guilders earned during the experiment plus a 10 Francs show-up fee.

The following pages explain the experiment in detail.

Brief Survey on the Experiment's Procedure

There are 8 employers and 16 employees. During the entire experiment you are an **employee**.

In each period of the experiment two employees are matched with one employer. The participants are, therefore, in each period in a group of three. After each period new groups of three are randomly formed. **In each period you are, therefore, in a group with new participants.** You will never be informed about the other group members' identity.

Each period consists of **two stages**:

Stage one

At the beginning of **stage one** the employer chooses a type of contract, which determines the wage level for the two employees. The employer determines both the wage of the employee with the higher **work output**, as well as the wage of the employee with the lower work output. The higher the employees' work outputs are, the higher is the employer's income.

Stage two

After the employer has made its decision on the type of contract, **at stage two** both employees decide on their **effort**. An employee's work output depends directly on the chosen effort. The work output depends, furthermore, on an element of luck. The employees have to bear costs when choosing effort

When both the employer and the employees have made their choice, the participants are informed about the resulting incomes. This ends a period.

There is a total of 12 periods.

The incomes from all periods are added up, converted into Francs and paid out in cash on completion of the experiment.

The Experiment's Procedure in detail

During the entire experiment you are an employee. In each of the 12 periods you are matched with another employer and another employee. The participants are divided into groups of three for the duration of each period. In each period new groups of three are randomly created. **Accordingly, in each of the twelve periods you are matched in a group with different participants.**

Stage one

In each period your employer chooses a contract, which determines the wage the employer pays to the employee with the **higher work output**, as well as the wage the employer pays to the employee with the **lower work output**. The following table shows the feasible types of contracts:

Type of Contract	Wage for the higher work output	Wage for the lower work output
1	150	150
2	160	140
3	170	130
4	180	120
5	190	110
6	200	100
7	210	90
8	220	80

As you can infer from the table the employer can either pay both employees the same wage (contract type 1) or it can pay them different wages (contract types 2 to 7). The wage sum is always the same, i.e., 300 Guilders.

As an example, the employer can choose to pay the employee with the higher work output a wage of 160 Guilders and the employee with the lower work output a wage of 140

Guilders (contract type 3). It may, however, also pay both employees a wage of 150 Guilders, regardless of their work output (contract type 1). Both you and the other employee hired by your employer, are informed about the contract type chosen by the employer.

Stage two

As soon as the employer has chosen a contract type, both employees choose an effort level. **The higher your effort, the higher your work output. This implies that the probability to get a higher wage is the higher, the higher the effort you have chosen.**

The work output, thus, depends on the effort chosen. In addition, your work output depends on a **random number**. This random number is a number between 0 and 10 randomly generated by the computer. This number is rounded up to four decimal places (e.g.,: 0.4212; 7.1703; 2.0000; ....). Each number (between 0 and 10) is drawn by the computer with the same probability.

Your work output is determined by your effort choice and the random number (between 0 and 10) generated by the computer:

Work output = effort + random number

Thus, your work output is the higher, the higher your effort. Moreover, your work output is the higher, the higher the random number. As an example, if you decided on an effort of 6 and the computer supplies a random number of 3.7900, your work output amounts to 9.7900, etc.

Choosing effort is associated with **costs**. In general, the higher your effort, the higher are the costs. The table below shows the **feasible efforts and the corresponding costs**:

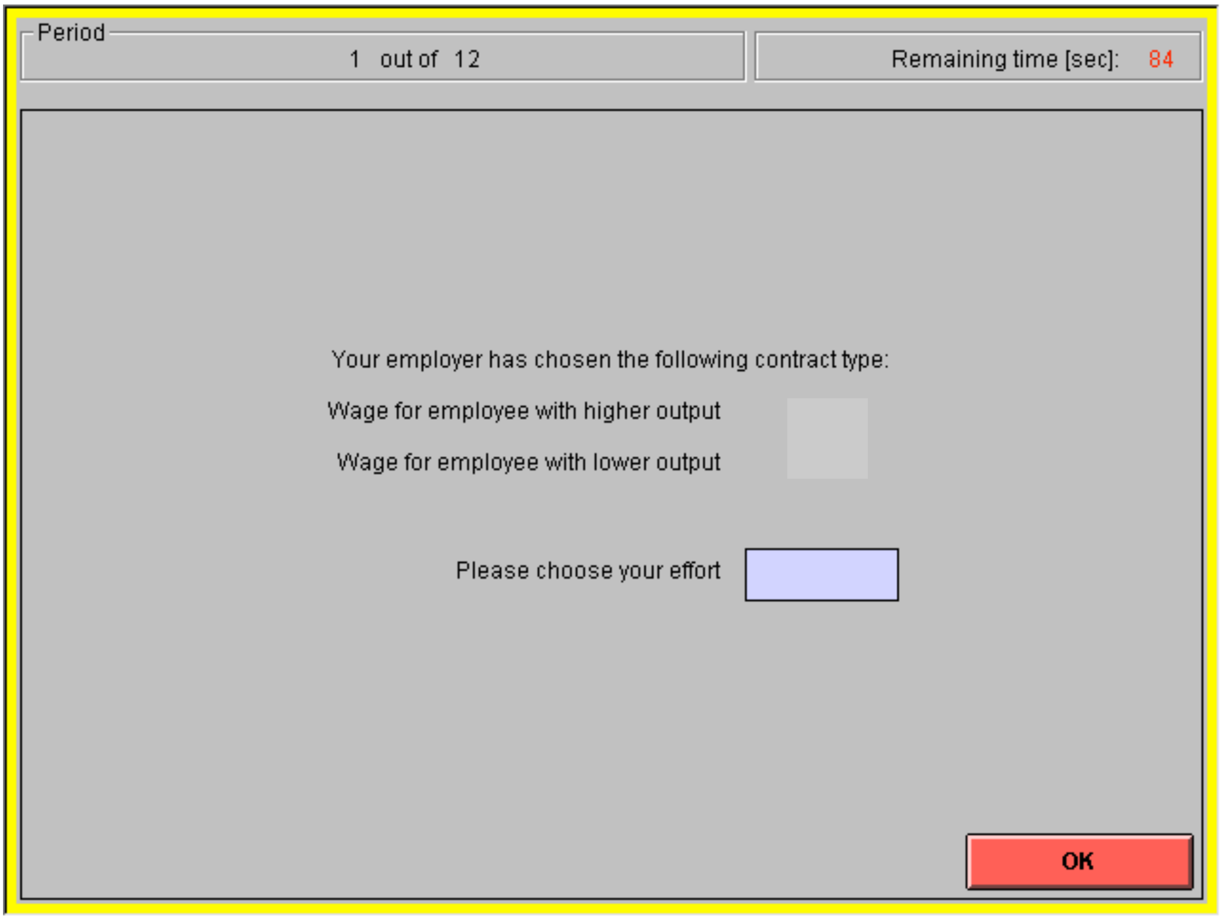
Feasible efforts and effort costs:

effort	**1**	**1.5**	**2**	**2.5**	**3**	**3.5**	**4**	**4.5**	**5**	**5.5**	**6**	**6.5**	**7**
effort cost	0	0.25	1	2.25	4	6.25	9	12.25	16	20.25	25	30.25	36

effort	**7.5**	**8**	**8.5**	**9**	**9.5**	**10**	**10.5**	**11**	**11.5**	**12**
effort cost	42.25	49	56.25	64	72.25	81	90.25	100	110.25	121

Feasible are all efforts between 1 and 12 in steps of 0,5.

At the beginning of each period the following **input screen** appears:



In the top left corner of the screen you see in which period you are. Below you see which type of contract "your" employer has chosen. You see which wage the employee with the higher work output gets and which wage the employee with the lower work output gets.

Below you see your input field. Please insert your effort here. Feasible are all efforts between 1; 1.5; 2; 2.5; ....; 12 (see table). Type the your effort choice in the corresponding input field and click the OK button. As long as the OK button is not clicked, you can revise your decision for this period.

How do you calculate your income?

If you achieved a higher work output than the other employee, you get the higher wage, otherwise you get the lower wage. Should by coincidence both achieve the same work output, it will be determined by chance who gets the higher and who gets the lower wage.

Let's assume the employer decided to pay the employee with the higher work output a wage of 160 Guilders and the employee with the lower work output a wage of 140 Guilders. If you achieve a work output of 7.1223 and the other employee achieves an output of 6.3711, your wage is 160 Guilders, and the other employee's wage is 140 Guilders.

From the received wage the effort costs are subtracted.

Your income is thus calculated as follows:

Your income = wage – effort costs

Thus, your income is the higher, the higher the wage and the lower your effort.

As an example, if you get a wage of 160 Guilders, and your effort costs are to 16 Guilders, your income amounts to 144 Guilders.

How to calculate the employee's income?

The employer's income depends on the work outputs of the two employees. The higher the efforts chosen by the employees and the higher their random numbers, the higher is the employer's income.

When the participants have made their decisions, **income screen** informs you about the **work outputs of both employees**, which **wage** you receive and your **income** in Guilders in the current period.

As soon as you click the "continue" button on the income screen, the current period of the experiment is concluded and the next period starts.

There is a total of 12 periods.

Are there any questions?

Control Exercises:

Please answer all questions and write down the whole calculation! If you have any questions please ask us!

1. You decided on an effort of 3.5, and the computer selected a random number of 1.7000.

   What is your work output: ....

2. The employer applied the following type of contract: The wage for the higher work output is 200, the wage for the lower work output is 100. You choose an effort of 6, and your random number is 7.0000. The other employee chooses an effort of 5 and gets a random number of 4.0000.

   What is your work output? ....

   What is the other employee's work output? ....

   What wage do you get? ....

   What wage does the other employee get? ....

3. You were paid a wage of 210, and your effort costs are 49.

   What is your income? ....

4. The employer pays the employee with the higher work output a wage of 190 Guilders, the employee with the lower work output a wage of 110 Guilders. Both employees decide on an effort of 7. Your random number is 3.1200, the random number of the other employee is 4.0000.

   What wage do you get? ....

5. The employer pays the employee with the higher work output a wage of 150 Guilders, the employee with the lower work output a wage of 150 Guilders. Both employees decide on an effort of 3. Your random number is 1.0000, the other employee's random number is 2.3400.

   What wage do you get? ....

6. Let us look at two cases: In case (a) both employees choose an effort of 3 and get a random number of 2.0000 and 0.1234 respectively. In case (b) both employees choose an effort of 8 and get a random number of 3.0000 and 6.1112 respectively.

   In which case is the employer's income higher? ....