Bonn Econ Discussion Papers

Discussion Paper 12/2004

Sabotage in Asymmetric Contests An Experimental Analysis

by

Christine Harbring, Bernd Irlenbusch, Matthias Kräkel, Reinhard Selten

July 2004

Bonn Graduate School of Economics
Department of Economics
University of Bonn
Adenauerallee 24 - 42
D-53113 Bonn

The Bonn Graduate School of Economics is sponsored by the

Deutsche Post World Net

MAIL EXPRESS LOGISTICS FINANCE

Sabotage in Asymmetric Contests – An Experimental Analysis

CHRISTINE HARBRING[#], BERND IRLENBUSCH[*], MATTHIAS KRÄKEL[♣], REINHARD SELTEN[◊]

Abstract

In a contest players compete for winning a prize by exerting effort and thereby increasing their probability of winning. Contestants, however, could also improve their own relative position by harming the other players. We experimentally analyze contests with heterogeneous and homogeneous agents who may individually sabotage each other. Our results suggest that sabotaging behavior systematically varies with the composition of different types of agents in a contest. Moreover, if the saboteur's identity is revealed sabotage decreases while retaliation motives prevail.

JEL Classification: C91, J33, J41, C72

Key Words: Contest, Experiments, Sabotage, Tournament

We thank Javier Sanchez Monzon who programmed the experimental software. Financial support by the *Deutsche Forschungsgemeinschaft* through grant KR2077/2-1 and SFB/TR 15 is gratefully acknowledged.

[#] University of Bonn, Department of Economics, BWL II

[*] University of Erfurt, Department of Economics, Microeconomics

[♣] Corresponding author: University of Bonn, Department of Economics, BWL II, Adenauerallee 24-42, D-53113 Bonn, Germany, tel.: ++49-228-739211, fax: ++49-228-739210, email: m.kraekel@uni-bonn.de

[◊] University of Bonn, Department of Economics, Laboratorium für experimentelle Wirtschaftsforschung

1 Introduction

Contests or rank-order tournaments describe competitive situations, in which players receive fixed prizes according to their relative performance. More successful players obtain higher prizes than less successful ones. In practice, there are a lot of examples for such contests: relative compensation of managers (Gibbons and Murphy 1990, Eriksson 1999), job promotion tournaments (Baker, Gibbs and Holmström 1994), contests between sales persons (Mantrala, Krafft and Weitz 2000), golf tournaments (Ehrenberg and Bognanno 1990a, 1990b), stock-car racing (Becker and Huselid 1992), horse racing (Lynch and Zax 1998) beside many others. Although contests are widely used as an incentive device, they suffer from a severe drawback. Since only the ordinal ranking of the agents' output is decisive for obtaining the winner prize, each agent could in principle improve his position not only by increasing his own output but also by exerting destructive effort to decrease the output of competitors, e.g. by bullying in job promotion tournaments. Actually, sabotaging is often less onerous than exerting constructive effort which makes sabotage to become a serious problem.

In our paper, we will focus on sabotage in contests with heterogeneous agents. Heterogeneity is modeled via the agents' cost functions, i.e. less able agents have a steeper cost function than more able ones. The combination of sabotage and heterogeneity is hardly considered in the literature, although it describes a highly realistic setting. For example, members of an organization typically have different abilities. While the subordinates themselves might be quite aware of the different characteristics of each other, e.g. because they have worked together for a certain time and may observe each other more accurately, their superiors often do not know the exact abilities. If the contestants can choose constructive effort as well as sabotage in this situation, different behavioral patterns may be possible. On the one hand, heterogeneity may destroy both constructive and destructive incentives (Lazear and Rosen 1981, pp. 861-863; McLaughlin 1988, pp. 246-247): The less able agents resign and drop out of the competition, since they are the presumable losers. Here resignation is optimal to save on effort costs. Due to their superiority, the more able agents will also decrease their effort levels as a best response. On the other hand, heterogeneity may increase the intensity of sabotage against certain contestants (Chen 2003). If there is only a small number of more able agents ("Favorites"), these Favorites will become visible as the most serious candidates for winning the contest. If, in addition, there are only few winner prizes (e.g., only one promotion to a better paid job), sabotage among the Favorites will be very intensive. Furthermore, if less able agents ("Underdogs") outnumber the competing Favorites, the Underdogs may (implicitly) form an alliance to concentrate sabotage on the Favorites. If, the other way round, many Favorites compete against a few Underdogs, sabotage between the Favorites may escalate in a “battle of the giants”. The investigation of these hypothetical behavioral patterns is at its very heart an empirical question.

In real-world contests, sabotaging can hardly be observed by a third party, because it is generally performed in secret.[1] Therefore, we have designed and conducted a controlled experiment in the laboratory to test the above hypotheses on resignation, superiority, visibility, alliance and battle of the giants. Moreover, we have distinguished two types of treatments. In half the treatments, the players do not learn the identity of the ones who have sabotaged them (no-information treatments). In the other half, the identity of the saboteurs is revealed (information treatments). This distinction has been made, since, in practice, victims of sabotage sometimes may be able to identify the culprits and sometimes not. For example, whether sabotaged members of an organization can identify the culprits may depend on the contenders’ closeness, their past experience, the knowledge of their colleagues’ characters and so on. To emphasize the importance of heterogeneity we also conduct symmetric treatments with homogeneous agents and contrast them with the asymmetric treatments.

Our experimental design is based on a logit-form contest.[2] This type of model has been used for modeling rent-seeking behavior (e.g., Tullock 1980), patent and innovation races (e.g., Baye and Hoppe 2003), sporting contests (Szymanski 2004), and labor market contests (Kräkel 2002).[3] The findings of this paper show that heterogeneity and the form of heterogeneity (i.e., the proportion of Favorites and Underdogs in a contest) is crucial for sabotage behavior in contests. For example, under both information conditions Underdogs sabotage Favorites significantly less the higher the proportion of Favorites. This result supports the resignation hypothesis and cannot be expected from a theoretic perspective. The experimental results also show that both the productive and the destructive behavior clearly differ between the no-information and the information settings. Our data provides persuasive evidence that behavior is guided by the desire to retaliate if the identity of saboteurs is revealed.

Only few papers theoretically address the problem of sabotage in contests. The first analysis has been done by Lazear (1989). His main result shows that the optimal spread between

---

[1] In real organizations sabotage such as refusing to cooperate with others, deliberately transferring false information, or refraining from passing viable information is usually forbidden. Data on real-world sabotage activities can therefore hardly be collected.

[2] For logit-form contests see, among many others, for example Tullock (1980), Dixit (1987), p. 893, Gradstein and Konrad (1999).

winner and loser prize is lower in tournaments with sabotage than in tournaments in which sabotaging is excluded. The intuition for this result comes from the fact that both constructive and destructive Nash equilibrium effort increases in the prize spread. Konrad (2000) considers sabotage behavior in logit-form contests. His analysis shows that sabotage is particularly serious in contests with a small number of contestants. If there are many contestants sabotage may become less attractive because sabotaging one player provides positive externalities to all other players. Chen (2003) uses a rank-order tournament framework similar to Lazear to analyze productive and destructive activities in tournaments when participants vary in their abilities. The underlying assumption is that contestants might be differently talented to sabotage other contestants and to increase their performance by increasing their productive effort. An interesting result of his work is that a participant who is more talented in terms of the effectiveness of his productive effort than the competitors is sabotaged by other players more often. Kräkel (2004) considers sabotaging – as well as helping – as an instrument to undermine competition in asymmetric tournaments. Similar to our paper, in a two-stage model players first decide on sabotaging/ helping and then have to choose their constructive efforts at the second stage. In this context, players have an incentive to use sabotage for preempting their competitors.

For obvious reasons empirical studies on sabotage in contests are quite rare. Garicano and Palacios-Huerta (2000) use field data from the Spanish national soccer league to test the implications of introducing the three-point rule in professional soccer. In the old system, the winner of a soccer match receives two points, but according to the new rule winning a match is rewarded by receiving three points. Hence, in terms of the contest theory the spread between winner and loser prize has been increased which, according to Lazear (1989), should result in higher productive and destructive efforts. The empirical findings confirm that the introduction of the three-point rule is indeed accompanied by significantly more constructive and destructive effort.[4] In another study, Drago and Garvey (1998) gather data based on questionnaires distributed among Australian companies and find that helping effort in workgroups is reduced when promotion incentives are large. Helping effort can be interpreted as the opposite behavior of sabotage and indicates the tendency of employees to cooperate with each other.

---

[3] From Loury (1979) we also know that a logit-form contest success function is equivalent to a rank-order tournament by Lazear and Rosen (1981) that is characterized by exponentially distributed noise and linear costs.

[4] Note that Garicano and Palacios-Huerta approximate the intensity of destructive activities by analyzing the number of defenders whose aim is to reduce the output of the competing team and the number of red/yellow cards. Thus, part of the sabotage activities analyzed here are legal.

There are several experimental papers on tournaments and contests that are related to our study, but they do not deal with sabotage among asymmetric agents.[5] These experiments examine either the impact of heterogeneity on incentives or the impact of sabotage in a setting with homogeneous agents. Tournaments with heterogeneous agents have been analyzed in several studies (see e.g., Bull, Schotter and Weigelt 1987, Weigelt, Dukerich, and Schotter 1989, Schotter and Weigelt 1992, and for a real-effort experiment see van Dijk, Sonnemans and van Winden 2001). The broad results of these investigations indicate that subjects that are disadvantaged ex ante either strain themselves all the more when competing against a more able player or drop out of the tournament, i.e. they exert a very low effort. Despite an eventual oversupply of the Underdog's effort the sorting of agents is still feasible as the stronger agent does not slack off and emanates from the tournament as the winner more often.[6] This finding is in line with the experiment of Müller and Schotter (2003) who analyze the influence of prizes in contests with heterogeneous agents following the theoretic work by Moldovanu and Sela (2001). Müller and Schotter present clear evidence that efforts of laboratory subjects are bipartite: While the low ability workers either drop out or exert only little effort the high ability workers oversupply effort.

The only experimental studies on sabotage in tournaments that we are aware of are Falk and Fehr (personal communication) and Harbring and Irlenbusch (2003a, b).[7] Their findings are based on tournaments with symmetric agents. They are consistent with those of Garicano and Palacios-Huerta (2000) as well as Drago and Garvey (1998) since both productive effort as well as the tendency to sabotage are significantly enhanced with an increased prize spread. We aim to fill a gap in the experimental contest literature by focusing on sabotage between heterogeneous players.

The paper is organized as follows. Section 2 introduces a two-stage contest model with sabotage which serves as the framework for our experiment. Section 3 summarizes the hypotheses we want to analyze. In section 4, our experimental design and procedure is described. The experimental results are presented in section 5. In section 6, we discuss three

---

5 In particular, for experiments on logit-form contests see for example Millner and Pratt (1991), Davis and Reilly (1998), Potters, de Vries, and van Winden (1998) and Weimann, Yang and Vogt (2002).

6 Harbring and Ruchala (2003) investigate the influence of an increase in the winner prize on behavior of heterogeneous agents. They, however, find that the prize difference has to be sufficiently high to identify Favorites.

7 In the work by Falk and Fehr sabotage is modeled as a binary choice between "sabotage" and "not sabotage" in a two-person tournament where sabotage always destroys the competitor's total output. Harbring and Irlenbusch analyze sabotage activities in a setting with more than two agents. The modeling of their sabotage activity allows for different levels of sabotage similar to the modeling of effort levels to be chosen.

additional treatments on the revelation of the saboteur's identity. Section 7 concludes the paper.

2 The Game

We consider a two-stage game with three players or contestants. At the first stage (*sabotage stage*), each player i chooses his intensity of sabotage s_{ij} against player j ($i, j = 1,2,3$; $i \neq j$). Sabotaging other players makes it more costly for them to exert effort. For simplicity, we concentrate on two sabotage levels s_L and s_H with $s_L < s_H$. After the first stage each player i observes S_i , the total amount of sabotage $\sum_{j\neq i} s_{ji}$ received from other players $j \neq i$ ($i = 1, 2, 3$). All total amounts of sabotage are revealed to all players resulting in the cost levels of each player. These cost levels represent the different marginal costs of effort that are due to the sabotage activities at the first stage. At the second stage (*contest stage*) the players choose their efforts e_i, while cost of effort depends on each player's cost level. Figure 1 summarizes the sequence of play:

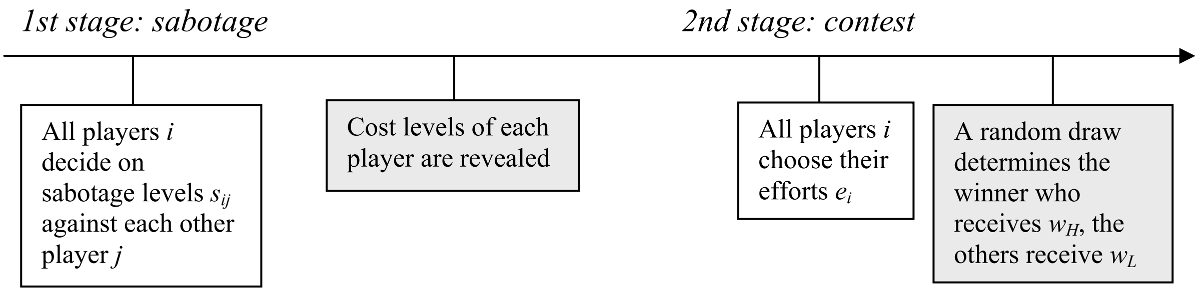


Figure 1: Sequence of the game

We adopt a contest success function of a logit-form contest which gives player i the winning probability $e_i /(e_i + \sum_{j\neq i} e_j)$ if $e_1 + e_2 + e_3 \neq 0$, and 1/3 otherwise. The contest winner receives the prize w_H, whereas each of the two other players obtains the loser prize w_L with $w_L < w_H$. Let $\Delta w = w_H - w_L$ denote the prize spread. Note that both activities – sabotaging and exerting effort – make it more difficult for the other players to win the contest. There is, however, an essential difference between the two activities: whereas sabotage hurts players individually, exerting effort simultaneously reduces the chance of both other players to receive the winner prize.

To keep the model manageable in the following analysis we assume that players are risk neutral and purely money maximizing. Thus, player *i*'s expected payoff function can be written as

$$E\Pi_i = w_L + \Delta w \frac{e_i}{e_i + \sum_{j \neq i} e_j} - k_i \cdot S_i \cdot e_i - c \cdot \left(\sum_{j \neq i} s_{ij} \right) \qquad (i = 1, 2, 3) \qquad (1)$$

with $k_i > 0$ and $c > 0$. In equation (1), the third term on the right-hand side describes player *i*'s costs of exerting constructive effort. The cost parameter k_i determines whether a player has ex ante low costs, i.e. $k_i = k_L$, and can be called a *Favorite*, or high costs, i.e. $k_i = k_H$ with $k_L < k_H$, and can be called an *Underdog*. As mentioned above S_i denotes the total amount of sabotage $\sum_{j \neq i} s_{ji}$ which player *i* receives. The fourth term in equation (1) denotes player *i*'s costs of sabotaging other players.

When looking for subgame perfect equilibria by backward induction we start by considering the contest stage. For each player *i* we obtain the first-order conditions:[8]

$$\Delta w \frac{\sum_{j \neq i} e_j}{\left(e_i + \sum_{j \neq i} e_j \right)^2} = k_i S_i \qquad (i = 1, 2, 3) \qquad (2)$$

Solving this system of equations for e_i gives:

$$e_i^* = 2\Delta w \frac{\left(\sum_{j \neq i} k_j S_j \right) - k_i S_i}{\left(k_i S_i + \sum_{j \neq i} k_j S_j \right)^2} \qquad (i = 1, 2, 3) \qquad (3)$$

Equation (3) shows that a player's equilibrium effort e_i^* decreases in his cost parameter k_i as well as in S_i the total amount of sabotage received. Hence, given equal amounts of received sabotage a Favorite always exerts more effort and has a higher probability of winning than an

[8] The second-order conditions are always satisfied.

Underdog. Moreover, the equilibrium effort depends on the marginal costs of effort of the other contestants.[9] Inserting (3) into the players' objective functions $E\Pi_i$ according to equation (1) leads to

$$E\Pi_i = w_L + \Delta w \frac{\left(\left(\sum_{j \neq i} k_j S_j\right) - k_i S_i\right)^2}{\left(k_i S_i + \sum_{j \neq i} k_j S_j\right)^2} - c\left(\sum_{j \neq i} s_{ij}\right) \qquad (i = 1, 2, 3) \qquad (4)$$

At the sabotage stage, the three players simultaneously choose $s_{ij} \in \{s_L, s_H\}$ $(i, j = 1, 2, 3;\ i \neq j)$ to maximize their expected utilities according to equation (4).[10]

3 Hypotheses

The aim of this research is to gain deeper insights into the behavior of heterogeneous agents in contests when they may harm each other. According to our model agents may sabotage each other at the first stage, i.e. increase their competitors' marginal costs of effort. Given these marginal costs of effort at the second stage, they compete for the prizes by exerting effort. As mentioned above there is a fundamental difference between the two activities: whereas players can individually hurt a specific other contestant by sabotaging, exerting effort simultaneously reduces the chance of both other players to receive the winner prize. Therefore, in a contest with heterogeneous agents a comparison of the individual sabotaging behavior from and against various types of players is particularly interesting. Thus, in our analysis we primarily focus on sabotaging behavior. In the following, we present our hypotheses.

---

[9] Note that equilibrium effort is not monotonous in $\sum_{j \neq i} k_j S_j$, the sum of others' marginal costs of effort. If $\sum_{j \neq i} k_j S_j$ is smaller than 3 $k_i S_i$, i.e. three times the own marginal cost of effort, the equilibrium effort is increasing in $\sum_{j \neq i} k_j S_j$. If $\sum_{j \neq i} k_j S_j$ is higher than 3 $k_i S_i$ the equilibrium effort decreases in $\sum_{j \neq i} k_j S_j$.

[10] Each player faces four alternatives at the sabotage stage: one can sabotage no other player, only one of the two other participants, or both other players. Consequently, in a given treatment, i.e. a given cost parameter combination (k_1, k_2, k_3), a 4x4x4 matrix describes all possible outcomes for the three players with each cell containing the expected payoffs (eq. 4) given the optimal effort choices following from the respective combination of sabotage activities (eq. 3). For all treatments reported in this paper one can derive from the matrices that in the subgame perfect equilibria agents choose full sabotage activities. The matrices will be provided by the authors upon request.

HYPOTHESIS "RESIGNATION":
Underdogs sabotage Favorites less the higher the proportion of Favorites.

Underdogs might feel that their chances of winning the contest is low if they have to compete against many superior Favorites. Thus, they might resign and withdraw their sabotage activity because it seems useless. As mentioned above other experiments on tournaments and contests indicate that subjects that are disadvantaged either drop out of the contest or overexert effort (e.g., Bull, Schotter, and Weigelt 1989, Müller and Schotter 2003). In our context, we conjecture that the tendency to withdraw sabotage increases with an increasing superiority of the competitors.

HYPOTHESIS "SUPERIORITY":
Favorites sabotage Underdogs less the higher the proportion of Favorites.

On the other hand, Favorites might consider the Underdogs' chance of winning to be quite low if they are competing against many Favorites. Thus, they might withdraw their sabotage activities against Underdogs because they feel themselves to be increasingly superior to Underdogs the more Favorites there are. Both hypotheses "RESIGNATION as well as "SUPERIORITY" are in line with theory (Lazear and Rosen 1981, pp. 861-863; McLaughlin 1988, pp. 246-247): The Underdogs, who are the presumable losers in the tournament, will choose low (constructive and destructive) efforts to save costs. In response the Favorites will also decrease their effort levels. Thus, tournaments will always lead to decreased incentives if it is common knowledge that agents have different abilities

HYPOTHESIS "VISIBILITY":
Favorites sabotage Favorites less the higher the proportion of Favorites.

We expect that competition among Favorites becomes stronger the fewer Favorites are in the contest. If there is only a small proportion of Favorites, these Favorites will become visible as the most serious candidates for winning the contest. Hence, sabotage activities among Favorites should be quite high. If there are e.g. only two Favorites, each of them might think that he has a very good chance of winning if he harms the other competing Favorite. Things might appear to be different if many or even only Favorites compete against each other and none of the Favorites feels particularly predestined to win the contest. In this case sabotage might appear to be much less effective.

HYPOTHESIS "ALLIANCE":
Underdogs sabotage Underdogs less than Favorites.

With this hypothesis we conjecture that Underdogs tend to sabotage Favorites more than other Underdogs. Favorites might be considered as the more serious competitors on the way to win the contest. Thus, Underdogs might tend to (implicitly) form an alliance against the Favorites. This hypothesis is in line with the theoretic result of Chen (2003) who finds that abler agents are being sabotaged more often.[11]

HYPOTHESIS "BATTLE OF THE GIANTS":
Favorites sabotage Favorites more than Underdogs.

The last hypothesis regarding the sabotaging behavior of agents refers to our conjecture that Favorites tend to concentrate on other Favorites when exerting sabotage. Again, other Favorites might be seen as the more serious opponents to succeed in the contest. Thus, they have to be harmed at the first place. Once more this hypothesis is in line with the result from Chen (2003).

HYPOTHESIS "EFFORT":

(i) Effort increases in the sum of the competitors' marginal cost of effort.

(ii) Effort decreases with the agent's own marginal cost of effort.

At the first stage of the game participants exert their sabotage activities, which determine the marginal costs of effort of each competitor. After the marginal costs of effort have been made common knowledge agents have to decide on the exertion of productive effort at the second stage. Given our theoretical analysis in the previous section we expect that the exertion of effort at the second stage depends on the contestants' marginal costs of effort. As we have seen above effort should decrease with the own marginal cost but increase in the competitors' marginal costs.[12]

Based on these conjectures, in the next section we describe our experimental setup, which is designed to test our hypotheses.

---

[11] The modeling of the contest in Chen (2003), however, differs from the context we use in this study. As illustrative examples for his finding Chen describes the situation that sometimes stars who seem to have good chances to win a contest in the beginning do not make it in the end. This might be due to heavy attacks, e.g. negative campaigning in presidential election contests aimed at a promising Favorite of the opposite party.

[12] Note, that as mentioned in footnote 9 equilibrium effort is not monotonously increasing in the others' marginal costs of effort. However, in the range we adopt in the experiment this is almost always the case.

4 Experimental Design and Procedure

The experiment was conducted in the *Laboratorium für experimentelle Wirtschaftsforschung* at the University of Bonn. All sessions were computerized and the software was developed by using the toolbox RatImage (Abbink and Sadrieh, 1995). In total 90 students of different disciplines were involved in the experiment. Every candidate was allowed to participate in one session only.

We implemented three treatments with three person contests and varied the heterogeneity of agents. For an overview see Table 1. In one treatment we implemented a *symmetric contests* in which all agents had identical ex-ante marginal costs, i.e. we have $k_i = k_H$ for each player i. We denote this treatment with *Hom*. In the two treatments with *asymmetric contests* there was one player i who had a different cost parameter $k_i \in \{k_L, k_H\}$ than the other two players. These treatments are denoted by *Hetx* with $x = U$ if the majority of contestants are Underdogs and $x = F$ if the majority are Favorites. The parameterization used in the experiment was as follows:[13] $w_L = 500$, $\Delta w = 1000$, s_L=1, s_H=1.4, k_L=1, k_H=1.4, c=1 and $e_i \in \{0, 1, 2, \cdots, 300\}$. Given these parameterization it is individually rational – in most cases even a dominant strategy – for each player i to choose the high sabotage level s_H against each other player j. Note that choosing low sabotage levels would be Pareto-efficient. A theoretical comparison of the symmetric and asymmetric contests yields that the average equilibrium effort is higher in the symmetric contest in which players are homogeneous.

Table 1: Experimental Design

	Type of players	# Subjects	# Observations	# Rounds
Hom	homogeneous (k_i = 1)	30	10	30
HetF	two Favorites (k_i = 1) and one Underdog (k_i = 1.4)	30	10	30
HetU	one Favorite (k_i = 1) and two Underdogs (k_i = 1.4)	30	10	30

We collected ten independent observations for each treatment. Each group of players consisted of three participants and the matching of the agents to groups was fixed for the whole experiment. To identify players over rounds a color (blue, red and yellow) was randomly assigned to each player. In the treatments with heterogeneity types were attributed to the players and they were informed by a note in their cubicles whether their type was a

[13] We deducted an amount of 5 Talers for the exertion of each sabotage activity, i.e. 10 Talers for sabotaging both competitors. Thus, the costs for sabotage were linearly transformed by $(12{,}5 \cdot \Sigma_{j \neq i}\, s_{ij}) - 25$. All payoffs and costs were given in "Talers" which is the fictitious laboratory currency during the experiment.

Favorite (low-cost type) or an Underdog (high-cost type). The cubicles were allocated by letting participants randomly draw cards.

Before starting the experiment the instructions were read aloud to all participants. The language was kept neutral, i.e. we did not use the term "sabotage" that had to be chosen but "a cost level for the other player". Moreover, the terms "contest" and "prize" were not mentioned. At the first stage (sabotage stage) subjects could decide to give the other player an "A" or a "B" that determined the other players' cost levels at stage two (contest stage). Participants were informed that receiving "AA" from the other players meant lowest cost for a certain effort number, obtaining cost level "AB" the middle cost level and "BB" the most expensive cost level. Choosing a "B" therefore meant sabotaging the other player by increasing his marginal cost of effort. Participants received cost tables depicting the cost for each number that could be chosen at the second stage, i.e. in a treatment with heterogeneous players subjects received six cost tables, three for each type of player showing all the three cost levels.[14] After all participants had made their decision at the first stage they entered the second stage.

Participants were informed about all players' cost levels before the second stage, i.e. they received information on the other players' decisions without knowing who decided for an "A" or a "B". Participants then decided on their effort, i.e. an integer number between 0 and 300. Costs of efforts were presented on the computer-screen and additionally could be traced in the cost tables. When all numbers were entered a wheel of fortune was shown on the screen which was partitioned according to each player's probability of winning the contest.[15] If the yellow player, for example, chose a number representing half of the sum of all numbers chosen in the group his yellow cake piece in the wheel of fortune covered 50 percent of the total wheel. The wheel of fortune then rotated and determined the winner who received w_H=1500 Talers. All participants in a group saw the same wheel of fortune and thus could identify the winner. The other two players received w_L=500 Talers. Payoffs of all players were displayed on the screen, and the next round began.

During the whole experiment participants were not allowed to communicate with each other. A session consisted of 30 repetitions of the same tournament setting. The sessions lasted for about 1.5 to 2.0 hours. During the experiment the payoffs were given in Talers and in the end

---

[14] A translation of the instructions and an example of a cost table are given in the Appendix. Original instructions were written in German. They are available upon request from the authors.

[15] For an example of the screenshot showing the wheel of fortune see the Appendix.

they were converted into Euro by a previously known exchange rate of 1 Euro per 1200 Talers. All subjects were paid anonymously.

5 Experimental results

This section reports our experimental findings especially with regard to our hypotheses derived above. We start our analysis with a first overview on aggregated behavior at both stages of the game and compare actual behavior with our theoretic prediction. We then focus on the observed sabotaging behavior. Following this investigation of behavior at the first stage we discuss the respective effort exerted at the contest stage.

5.1 Overview of results

Table 2 provides the average sabotage and effort choices as well as the theoretical predictions for behavior at both stages. Given the parameter values chosen in the experiment, rational, risk-neutral and purely money maximizing agents should always sabotage their competitors. The experimental data, however, draws a different picture. Participants exert substantially less sabotage than predicted by the theoretical analysis. Note, that a situation where all subjects choose to sabotage each other occurs in only 12% of all contests at maximum.[16]

Table 2: Overview of aggregated results

	Theoretic prediction	Experimental results		Theoretic prediction	Experimental results	
	Sabotage	Sabotage exerted	Sabotage received	Effort	Average effort	Frequency of Winning
Hom						
3 Favorites	100%	48.00%	48.00%	79.37	122.98	33.33%
HetF						
2 Favorites	100%	57.83%	55.33%	86.51	122.94	37.33%
1 Underdog	100%	49.50%	54.50%	37.07	80.79	25.33%
HetU						
1 Favorite	100%	49.83%	73.67%	89.04	126.97	43.33%
2 Underdogs	100%	70.25%	58.33%	49.47	89.04	28.33%

If we compare the average effort exerted in the experiment and the theoretic prediction regarding equilibrium behavior with maximal sabotage at the first stage it becomes evident

[16] In treatment *HetU* in 12% of all contests (30 rounds with 10 groups add up to a total of 300 contests in this treatment) all agents choose to sabotage every other person in the group. This figure amounts to 10.3% in *HetF* and 7.3% in *Hom*.

that the actual average effort is higher in all treatments (see also Table 2).[17] Note, however, that optimal behavior at the second stage is different if one considers the actually observed behavior at the first stage where subjects in general did not exert full sabotage. If one takes into account that participants choose these lower sabotage activities a different optimal average effort can be calculated which lies above the equilibrium effort as marginal cost of effort is lower.[18] Comparing the actual effort with this optimal effort conditioned on actual first stage choices, average efforts again are well above the theoretically predicted choices.[19] Furthermore, the actual frequency of winning the contest is in line with the conjecture that Favorites win the contest more often (Wilcoxon test: *HetF* and *HetU*, significant at 10% level, one-tailed).

Regarding the differences in behavior between the treatments we find two clear results. First of all, the average sabotage activities exerted are significantly higher in *HetU* than in *Hom* (Mann-Whitney U test, significant at 5% level, one-tailed). According to our theoretic prediction the average effort levels should be higher in the *Hom* treatment than in the *HetF* treatment with one Underdog, and effort in *HetF* should be higher than in *HetU* with two Underdogs. The only significant difference in effort behavior between treatments is that average effort is weakly significantly higher in *Hom* than in *HetU* (Mann-Whitney U test, significant at 10% level, one-tailed).

5.1 Sabotaging Behavior

Let us now analyze the sabotage activity more deeply. Figure 2 shows the development of sabotage frequencies over rounds for all three treatments. In all treatments sabotage activities appear to be increasing over the first ten rounds. In the following rounds 20 to 30 the frequencies are rather stable. The relation of sabotage frequencies between treatments mentioned above persists over all rounds: on average sabotage in *HetU* is higher than in *HetF*

---

17 If one compares the average effort of each independent observation with the theoretic prediction, effort lies significantly more often above the predicted level than below. This can be confirmed by the Binomial test with an event probability of 0.5 for *Hom* at a level of significance of 0.1%, for *HetF* (Favorites at 10%, the Underdog at 5%) and partly for *HetU* (Underdogs at 10%), all one-tailed.

18 The average optimal effort based on actual sabotage activities was calculated by weighting each average optimal effort regarding a certain constellation of marginal costs of effort, i.e. a certain constellation of sabotage activities and heterogeneity of players, with the frequency of this situation.

19 Applying the Binomial test average effort per independent observation lies significantly more often above the theoretic prediction than below in *Hom* and *HetU* at a significance level of 5% (one-tailed). In *HetF* effort per group is not significantly more often above the theoretic prediction although the average effort over all observations is considerably higher.

and is lowest in *Hom*. We still have to disentangle sabotaging behavior summarized in Table 2 regarding the actions of Underdogs and Favorites. Figure 3 shows the percentages of sabotage exerted *by* each type of agent *towards* each type of agent for all three treatments. Along this classification of exerted sabotage we are able to approach our hypotheses stated in section 3. Our first three hypotheses are based on a comparison of the sabotage behavior between the treatments. We compare the sabotage of Underdogs (Favorites) towards Favorites (Underdogs) between both asymmetric contests and the sabotage activities among Favorites between *HetF* and *Hom*.

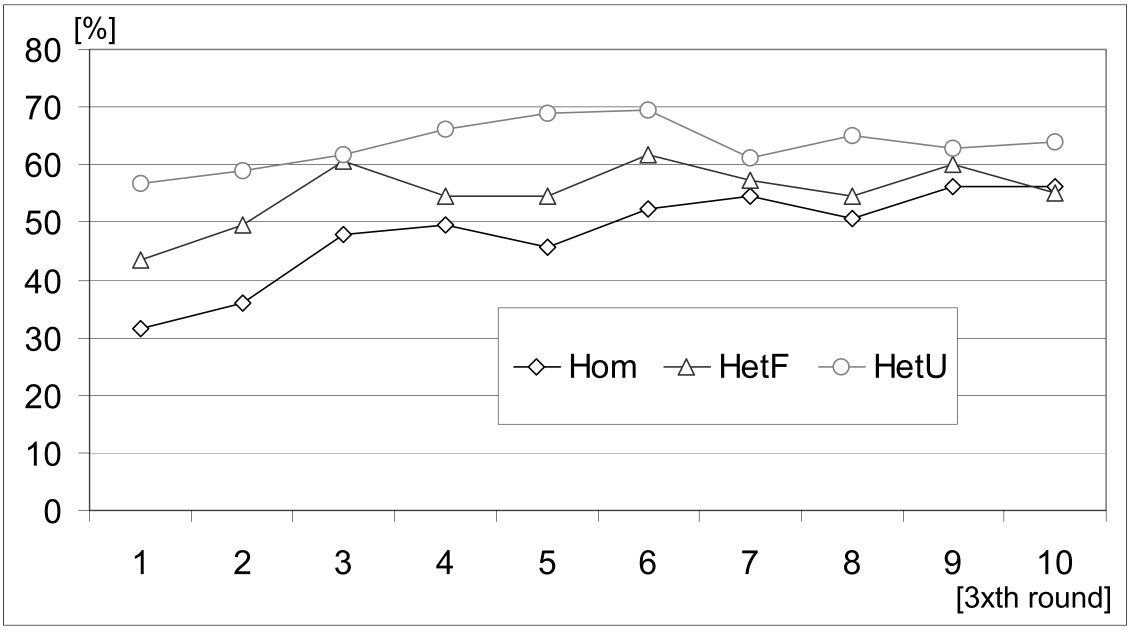


Figure 2: Percentage of sabotage over rounds (data points represent averages over 3 rounds)

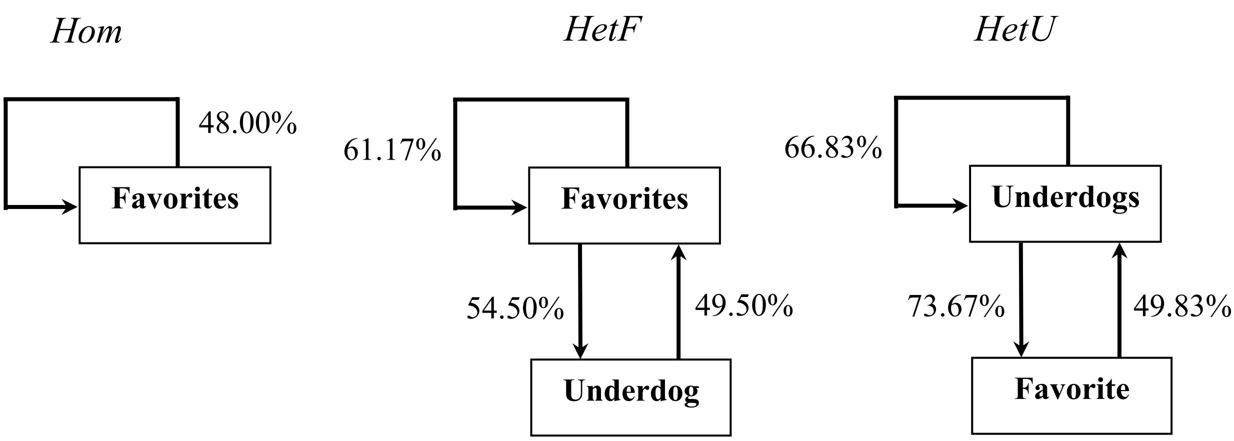


Figure 3: Individual sabotage exerted by each type of player

We conjecture that Underdogs sabotage Favorites less the higher the proportion of strong agents, i.e. Underdogs should sabotage Favorites less in *HetF* with two Favorites than in *HetU* with only one Favorite. Figure 3 already indicates that we find confirming evidence in favor of the hypothesis "RESIGNATION".

OBSERVATION ON HYPOTHESIS "RESIGNATION":
Underdogs sabotage Favorites less in *HetF* than in *HetU*.
(Mann-Whitney U test, significant at 5% level, one-tailed)

Thus, Underdogs seem to assume that their chance of winning decreases with an increasing proportion of Favorites and therefore exert less sabotage activity.

On the other hand, we conjecture that the Favorites sabotage Underdogs less the higher the proportion of Underdogs as they might feel superior such that high sabotage activities are not necessary. However, we cannot confirm the hypothesis "SUPERIORITY" as the sabotage activity of the Favorites aimed at Underdogs is not significantly larger in *HetU* than in *HetF*.

OBSERVATION ON HYPOTHESIS "SUPERIORITY":
Favorites *do not* sabotage Underdogs less in *HetF* than in *HetU*.

Our next hypothesis is based on a comparison between the Favorites' sabotage activities among themselves in the asymmetric treatment *HetF* with two Favorites and the symmetric contest with three Favorites. We can confirm our hypothesis "VISIBILITY".

OBSERVATION ON HYPOTHESIS "VISIBILITY":
Favorites sabotage Favorites less in *Hom* than in *HetF*.
(Mann-Whitney U test, significant at 5% level, one-tailed)

Interestingly, Favorites compete more fiercely against other Favorites with respect to sabotage activities if they face a lower number of these serious competitors.

In the following the observations on two further hypotheses are presented by focusing on the behavior of a specific type of agent within one treatment. We analyze the sabotage activities of Underdogs in *HetU* first. Underdogs in *HetU* may either sabotage the other Underdog or the Favorite. As stated above we conjecture that Underdogs tend to sabotage the Favorite, i.e. the "star" of the game rather than the other Underdog.

OBSERVATION ON HYPOTHESIS "ALLIANCE":
Underdogs *do not* sabotage Underdogs less than Favorites in *HetU*.

Although average activities seem to confirm our hypothesis the difference is not significant at a conventional level. However, we find that in total Favorites are targeted by sabotage than the Underdogs in *HetU* (Mann-Whitney U test, significant at 1% level, one-tailed, see also Table 2).[20]

Our last hypothesis on sabotage refers to the behavior of Favorites in *HetF* who may decide on sabotage towards the other Favorite and the Underdog. We conjecture that Favorites tend to sabotage other Favorites more often than the Underdog.

OBSERVATION ON HYPOTHESIS "BATTLE OF THE GIANTS":
Favorites sabotage each other more than the Underdog in *HetF*.
(Mann-Whitney U test, weakly significant at 10% level, one-tailed)

This seems to be in line with our observation on the hypothesis "VISIBILITY" from above. Favorites seem to concentrate on their "giants' battle".

5.2 Effort at the Contest Stage

The average effort exerted over rounds in the three treatments is shown in Figure 4. In each treatment average effort keeps relatively constant over rounds. While average effort in *Hom* is almost always higher than in the heterogeneous treatments, the highest of the latter two alternates over rounds.

[20] This is also due to the fact that Underdogs sabotage the Favorite slightly more often than the Favorite the Underdogs (Mann-Whitney U test, significant at 10% level, one-tailed).

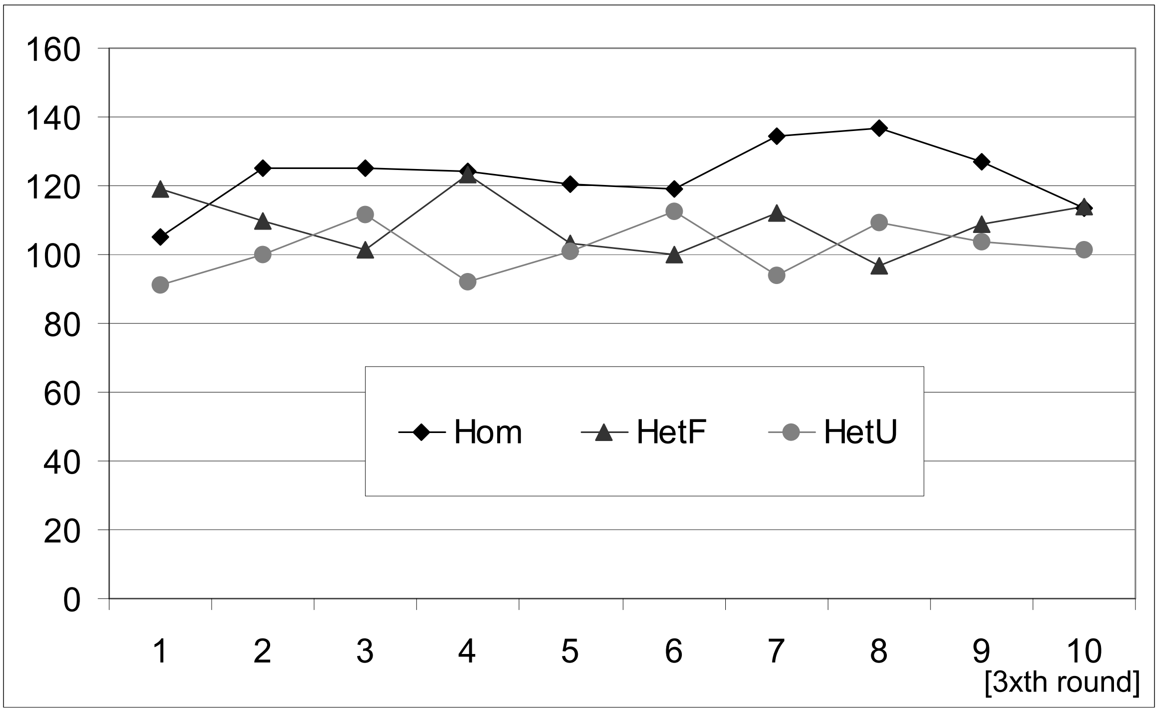


Figure 4: Average effort over rounds (data points represent averages over 3 rounds)

According to the hypothesis "EFFORT" the exertion of effort at the second stage depends on the marginal costs of effort of contestants which is qualitatively in line with our theoretic prediction. In section 3, we mentioned already that the exertion of effort depends on the own marginal cost of effort as well as the sum of marginal costs of the competitors. Note, that a competitor's marginal cost is a compound of the heterogeneity parameter as well as the amount of sabotage received by others. Table 3 denotes those marginal costs of effort which a participant might be facing. Every player can sabotage both other players in the three-person contest, i.e. there exist three feasible levels of being sabotaged: C_0 denotes the cost level when no other player exerts sabotage on that player, C_1 is the cost level when one is sabotaged by one other player, and C_2 means that one is sabotaged by both other players.

Table 3: Feasible marginal costs of effort $k_i S_i$

cost type	low cost type (Favorite)			high cost type (Underdog)		
sabotage received	C_0	C_1	C_2	C_0	C_1	C_2
Hom	2	2.4	2.8	-	-	-
HetF, *HetU*	2	2.4	2.8	2.8	3.36	3.92

Figure 5 depicts the effort levels predicted by theory as well as the actually observed average effort levels dependent on the constellation of the agent's own marginal cost of effort and the other contestants' marginal costs. Visual inspection suggests that average observed effort choices are qualitatively in line with theory.

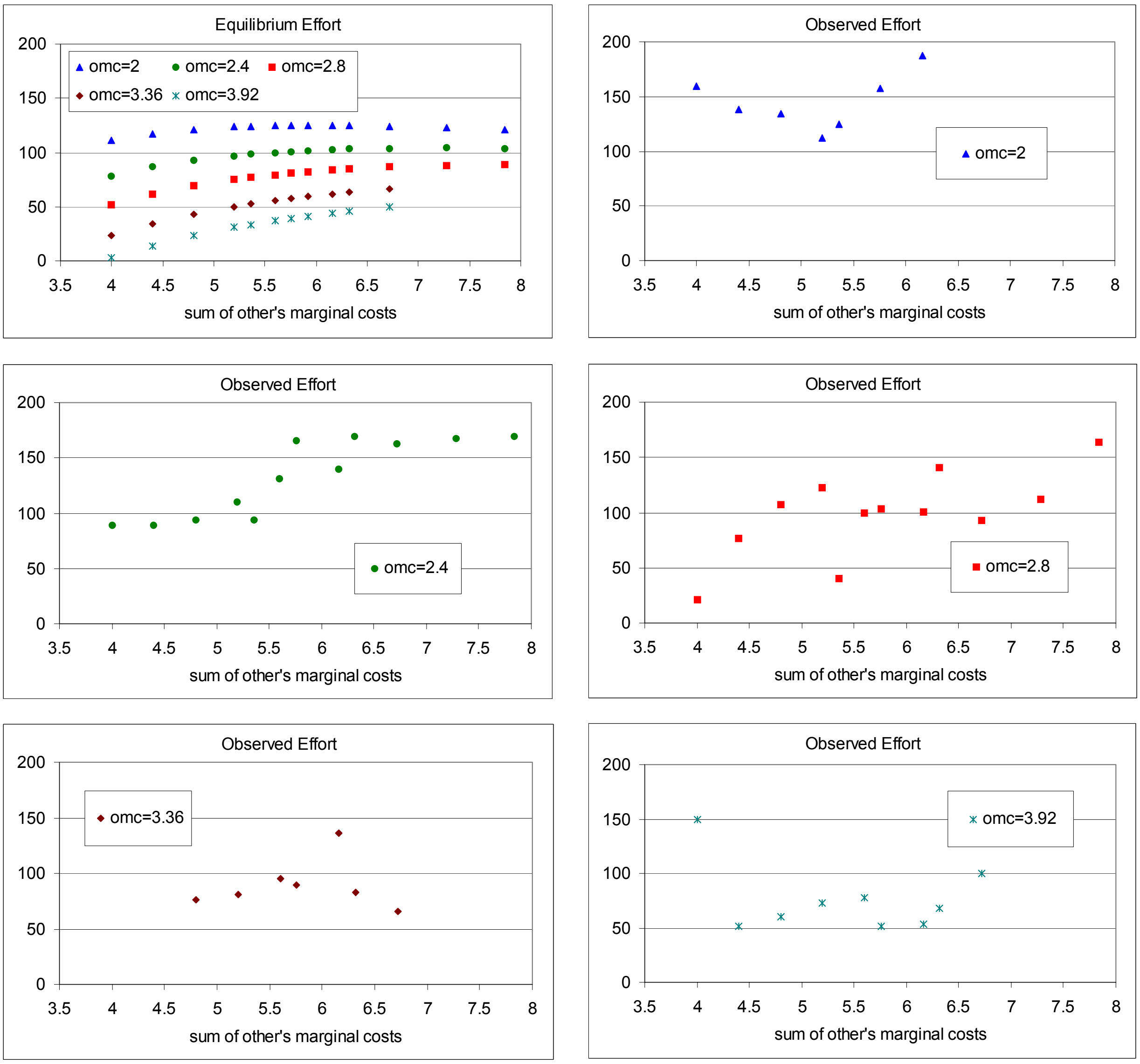


Figure 5: Theoretically predicted and actually observed[21] effort levels dependent on the constellation of marginal costs of effort (omc = own marginal cost of effort)

The statistical analysis yields that effort is significantly increasing with the sum of the competitors' marginal cost of effort.[22] Moreover, the effort level chosen is also in line with the theoretic prediction regarding the influence of the own marginal cost of effort.[23]

---

[21] Average efforts that occur in less than three groups or less than ten times are not depicted in the figure.

[22] We analyze the trend of average effort choices over the sum of marginal costs of competitors while keeping the own marginal cost of effort constant, i.e., the Spearman rank correlation coefficient between these average efforts and the sum of marginal costs of the competitors is computed for each own marginal cost given. By applying the Binomial test with an event probability of 0.5 we find that the correlation coefficients are significantly more often positive than negative at a significance level of 5% (one-tailed).

OBSERVATION ON HYPOTHESIS "EFFORT":

Effort increases with the sum of the competitors' marginal cost of effort.

Effort decreases with the agent's own marginal cost of effort.

The data reveals one interesting behavioral pattern regarding the exertion of effort that can be observed in some groups and is not expected from the theoretic prediction. Obviously, some players tend to increase their effort with the amount of sabotage they receive. This is in line with the behavior of several Underdogs who choose an effort level that causes them costs which even exceed the prize difference of 1000 Taler when being maximally sabotaged in the stage before. Thus, in total these players would be been better off by not exerting any effort at all. This phenomenon of such "more than ever"-types of players, however, is not a representative behavioral pattern.[24]

We designed our experiment such that both the cost parameter that introduces heterogeneity among agents and the sabotage parameter are chosen from the same set, i.e. $k_i \in \{1,\ 1.4\}$ and $s_{ij} \in \{1,\ 1.4\}$. Therefore, identical marginal costs $k_i\ S_i$ may result from different constellations of the cost parameters k_i and the amount of received sabotage S_i (see Table 3). Let us consider the constellations $k_1S_1 - k_2S_2 - k_3S_3$, i.e. all combinations of marginal costs for the three players. One can show that there are 10 different constellations in the symmetric contest. In the asymmetric contests there are 18 different constellations. Note, that several constellations might occur in the symmetric contests as well as in the asymmetric ones, e.g. 2.8 – 2.8 – 2.8, 2.0 – 2.8 – 2.8, 2.0 – 2.0 – 2.8, 2.4 – 2.8 – 2.8, 2.4 – 2.4 – 2.8.

As effort only depends on the constellation of marginal costs of effort, i.e. the own marginal cost of effort and the competitors' marginal costs according to theory, subjects should behave identically if the constellation of marginal costs of effort is identical. Thus, their behavior should not be influenced by the way the marginal costs are generated, by sabotage or by ex ante heterogeneity. Figure 6 depicts the average effort in two treatments regarding a specific

---

[23] The difference between the average effort for two adjacent own marginal costs of effort is computed while keeping the sum of the marginal costs of competitors constant. We count how often this difference is positive. By applying the Binomial test we find that more than half of the differences are significantly more often positive than negative at a level of significance of at least 5% (one-tailed).

[24] If participants decide to retaliate for being sabotaged they can actually opt for two alternatives: first, by sabotaging other players in the following round, or second, by exerting high efforts and hereby reducing the probability of winning for the other players. If one takes a look at the distribution of effort costs one observes that there are always agents being sabotaged twice who exert maximal effort and thus, have to bear maximal cost of effort. This behavioral pattern is quite consistent in all treatments and supports the conjecture that participants retaliate for a high cost level by exerting maximal effort and decreasing other agents' winning probability.

cost constellation: 2.4 – 2.4 – 2.8.[25] Efforts are averaged over the two players who face a marginal cost of 2.4. The sabotage received as well as the type of player is depicted. Taking a look at the amount of average efforts already reveals an interesting finding. Obviously, the emergence of the marginal cost of effort (ex ante heterogeneity or sabotage) affects the behavior of players.

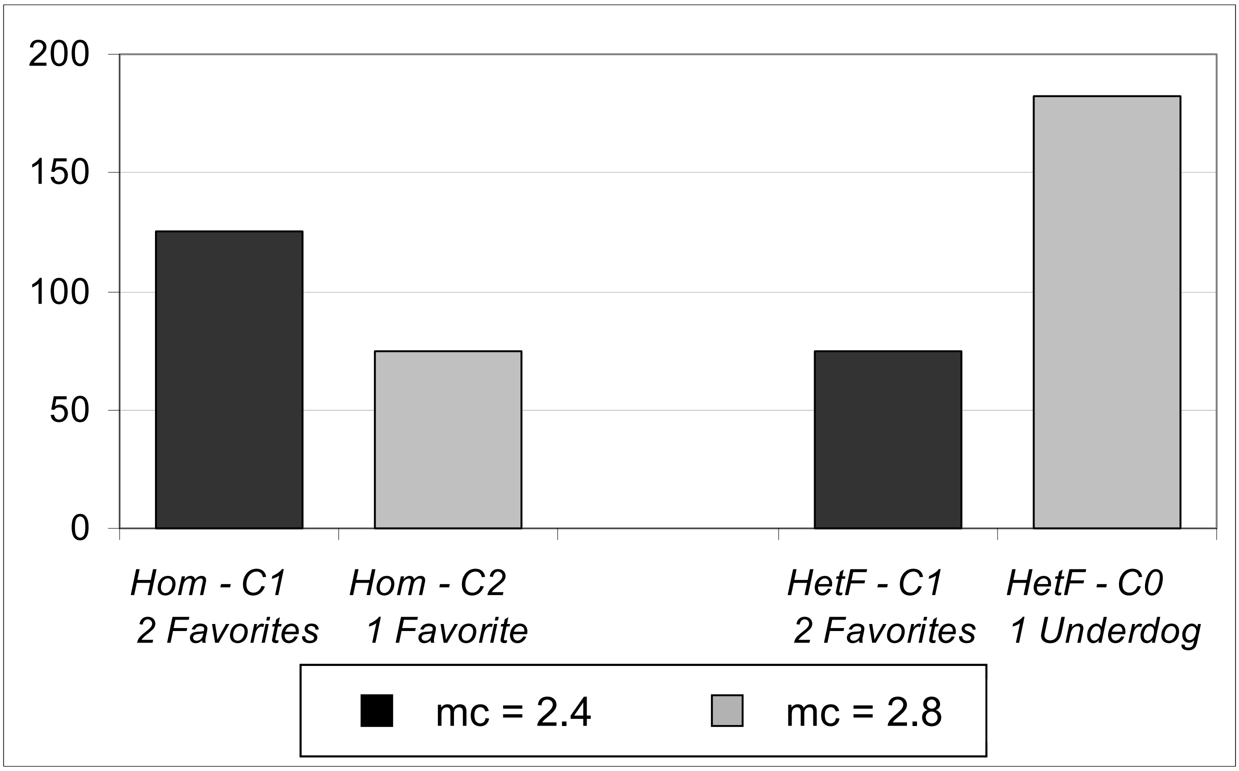


Figure 6: Average effort in different treatments with the same constellation of marginal costs (mc)

In the homogenous treatment average efforts are qualitatively in line with our theoretic prediction. The subject, who faces the higher marginal cost exerts a significantly lower effort than the one with a lower marginal cost (Wilcoxon test: *Hom* C_1 vs. *Hom* C_2 α = .05, one-tailed). In the asymmetric treatment, however, a different behavioral pattern emerges (see Figure 6). Surprisingly, in *HetF* the participants facing the higher marginal cost of 2.8 exert an even higher (weakly significant) average effort than the subjects with lower marginal cost (Wilcoxon test: *HetF* [marginal cost = 2.4: Favorite and C_1] vs. *HetF* [marginal cost = 2.8: Underdog and C_0] α = .1, two-tailed). If one compares the average efforts of those subjects with the same marginal costs between treatments one can see that the average efforts of those

[25] There are still other constellations, which are identical in different treatments, but not all constellations occur often enough to compare them in a statistically meaningful way.

players with lower marginal cost are significantly higher in the homogenous treatment than in the heterogeneous one (Mann-Whitney U test: *Hom* C_1 vs. *HetF* [marginal cost = 2.4: Favorite and C_1] α=.02, two-tailed).

Thus, it seems that the way a cost factor comes into existence plays an important role for the behavior of participants. The player with higher marginal cost in the above constellation in the asymmetric treatment is the Underdog on a cost level of C_0. Note, that this player suffers from an ex ante disadvantage due to his type, but nevertheless, his marginal costs are the lowest ones he can ever obtain, i.e. he did not experience any sabotage by another player. An interesting behavioral motive might be that being intentionally sabotaged by others discourages subjects more than being discriminated ex ante.

6 Revelation of Saboteurs' Identity

In the treatments analyzed so far it is not revealed who is sabotaged by whom. Thus, an agent is not able to identify the saboteur if he is sabotaged by only one other player. Note, that an agent knows the identities of saboteurs if he receives sabotage of both competitors. Although from a theoretical point of view the revelation of the saboteurs' identity should not be decisive for sabotaging behavior, it appears to be natural to assume that in a repeated setting there might be a considerable influence.

First of all, if one's identity is revealed, a participant who experienced a well-directed sabotage activity is likely to retaliate this unfriendly behavior in the next round by sabotaging the culprit herself.

HYPOTHESIS "RETALIATION":
Sabotage is retaliated by sabotage.

This is in line with our next hypothesis. We conjecture that as long as agents stay anonymous when making the contest more onerous for other participants, one might lose one's qualms and be rather tempted to exert higher sabotage.

HYPOTHESIS "REVELATION OF SABOTEURS":
Sabotage is lower if the saboteur's identity is revealed.

In order to test these two hypotheses we analyze three additional treatments, denoted by *HomInfo*, *HetFInfo* and *HetUInfo*, which are conducted with additional 90 participants. These treatments are identical to those investigated so far except for one feature: Now, the identities of saboteurs are *revealed* after the first stage. In the next subsection, these three *Info* treatments are compared to the *NoInfo* treatments we have analyzed in the preceding sections. Following this, we investigate the sabotage activities in the *Info* settings regarding our hypotheses introduced in section 3.

6.1 The influence of information

As illustrated by Figure 7 we can confirm our first hypothesis regarding the retaliation of sabotage activities in the *Info* settings.

OBSERVATION ON HYPOTHESIS "RETALIATION":

In the *Info* treatments an agent *X* more frequently sabotages an agent *Y* if that agent *Y* has sabotaged the agent *X* in the preceding round than if he had not.
(Wilcoxon-Signed-Rank test , *HomInfo*: significant at 10 % level, *HetFInfo* and *HetUInfo*: significant at 1% level, all one-tailed)

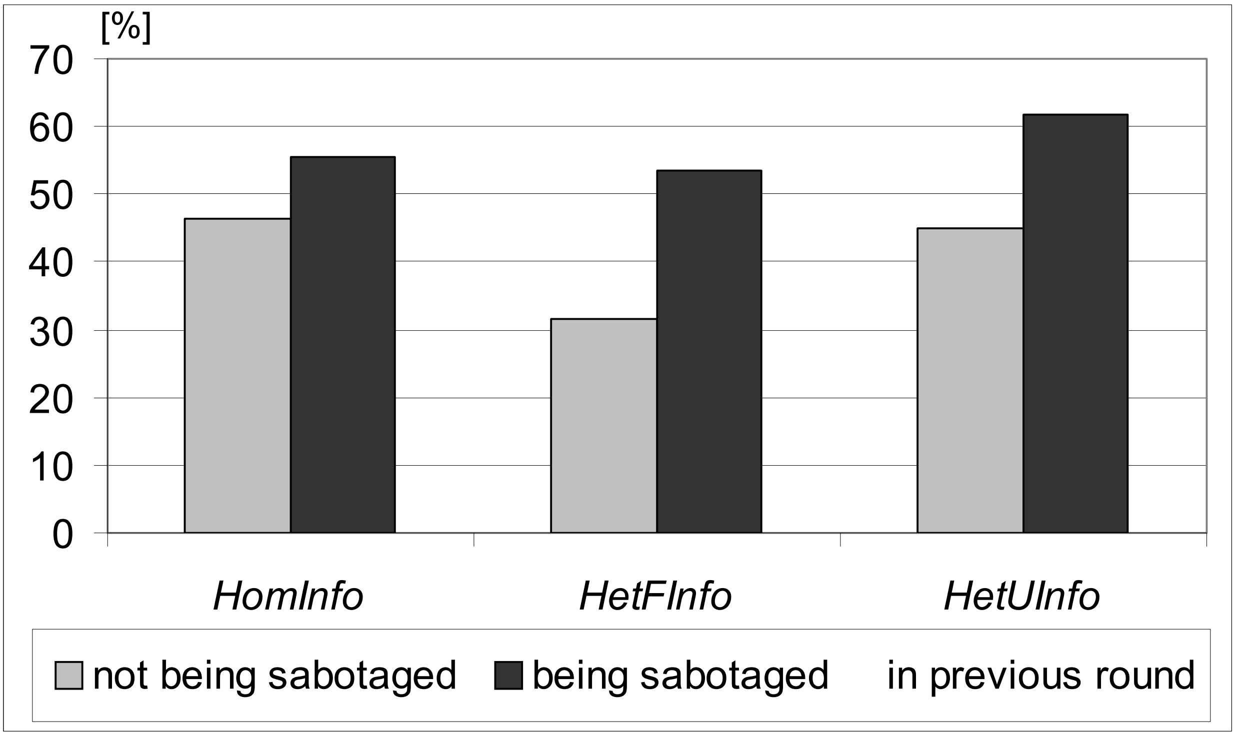


Figure 7: Frequency of sabotaging a contestant dependent on being sabotaged from the same one in the previous round

Our second hypothesis can only be confirmed for contests with heterogeneous participants:

OBSERVATION ON HYPOTHESIS "REVELATION OF SABOTEURS":

Sabotage in *HetFInfo* (*HetUInfo*) is lower than in *HetF* (*HetU*).

(Mann-Whitney U test, *HetF* vs. *HetFInfo*: 5% level; *HetU* vs. *HetUInfo*: 10% level, one-tailed)

This observation is based on aggregated behavior of both types of agents. To disentangle the sabotage activities of Favorites and Underdogs Table 4 denotes the percentage of sabotage exerted by each type of player. Interestingly, only the two Favorites in *HetFInfo* sabotage significantly less than in *HetF* (Mann-Whitney U test, significant at 5% level, one-tailed), and the two Underdogs in *HetUInfo* exert significantly less sabotage than in *HetU* (Mann-Whitney U test, significant at 5% level, one-tailed). Hence, the group of players who are of the same type in each treatment seems to be responsible for the reduction in sabotage exertion if additional information is supplied. Table 4 also provides an overview over the frequencies of sabotage received by each type of player. Analyzing the experimental results by pairwise comparison of the treatments yields that the Underdog in *HetF* suffers significantly more often from sabotage than in *HetFInfo* (Mann-Whitney U test: significant at 5% level, one-tailed), and in *HetU* the Favorite suffers from being sabotaged significantly more often than in *HetUInfo* (Mann-Whitney U test: significant at 5% level, one-tailed).

Table 4: Sabotage per type of player in *Info* and *NoInfo* treatments

	Sabotage exerted		Sabotage received	
	Info	*NoInfo*	*Info*	*NoInfo*
Hom				
3 Favorites	54.44%	48.00%	54.44%	48.00%
HetF				
2 Favorites	39.08%	57.83%	42.50%	55.33%
1 Underdog	43.17%	49.50%	36.33%	54.50%
HetU				
1 Favorite	60.67%	49.83%	55.67%	73.67%
2 Underdogs	48.75%	70.25%	51.25%	58.33%

However, Table 4 does not indicate whether sabotage towards both competitors is decreased in the *Info* settings. Table 5 summarizes the results regarding the sabotage activity exerted by each type of player as well as which type of player suffers from it. It is quite obvious that the group of players who are of the same type in each of the asymmetric treatments exerts higher sabotage activities towards both types of players under the *NoInfo* condition than under the *Info* condition.[26]

Table 5: Sabotage per type of agent and type of agent aimed at

	Underdog sabotages Favorite		Favorite sabotages Favorite		Underdog sabotages Underdog		Favorite sabotages Underdog	
	Info	*NoInfo*	*Info*	*NoInfo*	*Info*	*NoInfo*	*Info*	*NoInfo*
Hom								
3 Favorites			54.44%	48.00%				
HetF								
1 Underdog	43.17%	49.50%						
2 Favorites			41.83%	<** 61.17%			36.33%	<** 54.50%
HetU								
2 Underdogs	55.67%	<** 73.67%			41.83%	<** 66.83%		
1 Favorite							60.67%	49.83%

By using the Mann-Whitney U test (one-tailed) the amount of sabotage exerted per statistically independent observation can be compared between the treatments. The level of significance at which the null hypothesis can be rejected in favor of the directed alternative hypothesis is denoted by:

*	*weakly significant:*	$5\% < \alpha \leq 10\%$
**	*significant:*	$1\% < \alpha \leq 5\%$
***	*highly significant:*	$\alpha \leq 1\%$

As mentioned above an additional piece of information in the *Info* treatments compared to the *NoInfo* treatments is only supplied if an agent has been sabotaged by *one* other player in the group. In both other cases – when he has been sabotaged by none or both of the competitors – each player receives the same information in both treatment variations because the cost levels are revealed to them. Thus, it appears to be interesting to investigate the agents' reactions to being sabotaged once in both settings. Table 6 denotes how often a player sabotages one other player after being sabotaged once in the preceding round and how often a player sabotages both other players as a reaction to being sabotaged by one player. It is obvious that the type of player that is represented twice in each group significantly more frequently sabotages both competitors in the *NoInfo* condition than in the *Info* condition.[27] Moreover, in the asymmetric treatments the tendency to sabotage only one other player as a reaction to being sabotaged once is higher in the *Info* treatments.

---

[26] In each round players may opt for sabotaging none, one or both of their competitors. If the amount of sabotage each subject exerts in each round is analyzed we observe that the reduction of the sabotage activity in the *Info* conditions compared to the *NoInfo* treatments is mainly due to an increased frequency of rounds where no other player is sabotaged and a decrease of the frequency of rounds where both other players in the group are sabotaged.

[27] Only the type of player that is represented twice in each asymmetric treatment sabotages significantly more often both other players after being sabotaged once in the *NoInfo* conditions than one other player (Wilcoxon test, Favorites in *HetF* and Underdogs in *HetU*: significant at 5% level, one-tailed).

Table 6: Reaction via sabotage of agents to being sabotaged in preceding round

	Sabotage of **one** player after being sabotaged once in preceding round			Sabotage of **both** players after being sabotaged once in preceding round		
	Info		*NoInfo*	*Info*		*NoInfo*
Hom						
3 Favorites	33.08%		36.29%	37.40%		28.44%
HetF						
1 Underdog	26.72%		30.31%	36.38%		32.90%
2 Favorites	40.08%	>**	24.66%	21.16%	<***	46.16%
HetU						
2 Underdogs	35.30%	>*	27.19%	33.37%	<*	58.45%
1 Favorite	39.17%	>*	20.59%	37.16%		40.75%

By using the Mann-Whitney U test (one-tailed) the amount of sabotage exerted per statistically independent observation can be compared between the treatments. The level of significance at which the null hypothesis can be rejected in favor of the directed alternative hypothesis is denoted by:

*	*weakly significant:*	$5\% < \alpha \leq 10\%$
**	*significant:*	$1\% < \alpha \leq 5\%$
***	*highly significant:*	$\alpha \leq 1\%$

This finding might be explained by an aversion of subjects to treat equal types of players differently. In *HetFInfo* for example, an Underdog who has been sabotaged by one other player rather tends to retaliate this behavior by sabotaging both Favorites and not only the Favorite who had actually harmed him. Thus, the additional piece of information on the identity of the saboteur is not relevant for his act of retaliation. This applies analogically for the Favorite in *HetUInfo*. A Favorite in *HetFInfo* or an Underdog in *HetUInfo*, however, is confronted with a sabotage activity of one of the different types of players. If he knows the identity of the culprit only this type of agent is sabotaged. If he is not given the information on the identity in the *NoInfo* settings he tends to sabotage both competitors.

6.2 Sabotage activities in the *Info* treatments

This effect that the sabotage activities of those types of agents who are represented twice in the asymmetric treatments is decreased by supplying the information on the identity of the saboteurs affects the observations on the hypotheses from section 3. Figure 8 depicts the individual sabotage activities by each type of player.

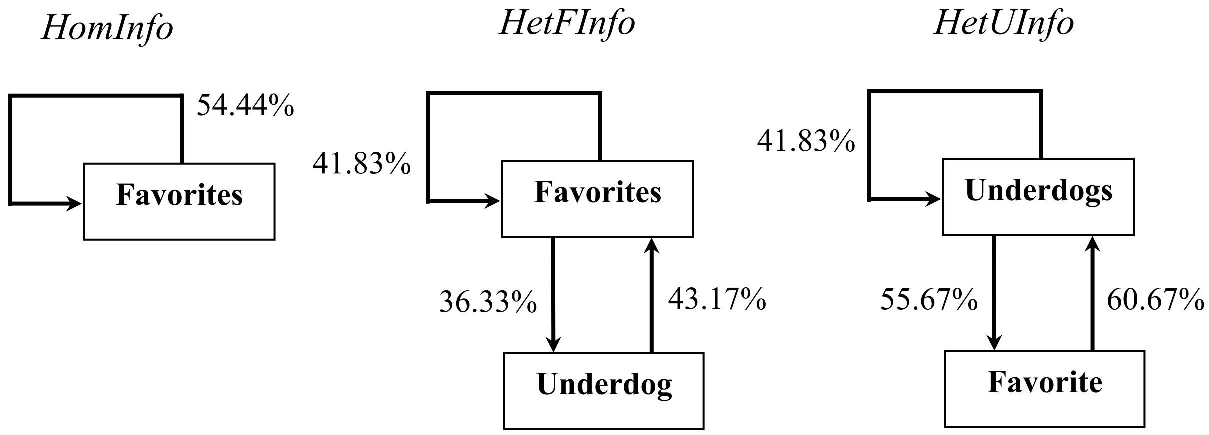


Figure 8: Individual sabotage exerted by each type of player in the *Info* settings

OBSERVATION ON HYPOTHESIS "RESIGNATION":
Underdogs sabotage Favorites less in *HetFInfo* than in *HetUInfo*.
(Mann-Whitney U test, significant at 10% level, one-tailed)

We can still weakly confirm our first hypothesis. However, this observation is weakened by the fact that Underdogs in *HetUInfo* reduce their sabotage activities if the culprits' identities are revealed.

Analogically, the Favorites exert higher sabotage activities in *HetF* with no information given on the identity of saboteurs than in *HetFInfo* with information supplied (see also Tables 4 and 5). While we cannot confirm our hypothesis "Superiority" for the *NoInfo* settings we can now confirm the hypothesis for the *Info* treatments:

OBSERVATION ON HYPOTHESIS "SUPERIORITY":
Favorites sabotage Underdogs less in *HetFInfo* than in *HetUInfo*.
(Mann-Whitney U test, significant at 5% level, one-tailed)

The fact that the Favorites exert lower sabotage activities in *HetFInfo* is also reflected by the observation on our next hypothesis.

OBSERVATION ON HYPOTHESIS "VISIBILITY":
Favorites *do not* sabotage Favorites less in *HomInfo* than in *HetFInfo*.

The observations on the next two hypotheses do not directly follow from our results regarding the influence of information from above. The Underdogs reduce their sabotage in *HetU* if information on saboteurs' identities is supplied but it is not obvious whether Underdogs tend to sabotage both of their competitors less.

OBSERVATION ON HYPOTHESIS "ALLIANCE":

Underdogs sabotage Underdogs less than Favorites in *HetUInfo*.
(Mann-Whitney U test, significant at 5% level, one-tailed)

The observation on our hypothesis "Alliance" weakly indicates that Underdogs tend to decrease their sabotage activities more vigorously towards the other Underdog than towards the competing Favorite.

This uncertainty about the reduction of sabotage applies also to our last hypothesis regarding the activities of the Favorites in *HetFInfo*.

OBSERVATION ON HYPOTHESIS "BATTLE OF THE GIANTS":

Favorites *do not* sabotage Favorites more than Underdogs in *HetFInfo*.

Thus, the observation on our last two hypotheses suggest that the type of player that is represented twice in the asymmetric contests decreases the sabotage activities rather towards the same type of player than towards the other competing type if the identity of the saboteur's identity is revealed.

7 Conclusion

This paper reports on the first experimental study to investigate the behavior of heterogeneous agents in contests with sabotage. At the first stage of the game three contestants decide on sabotage by a binary choice for each of the other players individually. By sabotaging the marginal cost of effort of the sabotaged agent is increased. At the second stage all participants choose their effort. The relation of effort choices determines the probability of obtaining the high winner prize. A special focus of our analysis is the heterogeneity among agents. We vary agents with regard to ex ante marginal cost of effort, denoting low cost types as "Favorites" and high cost types as "Underdogs". Thus, ex post marginal cost of effort is a compound of the ex ante type of agent and the amount of sabotage received in each round by both competitors.

Contrary to our theoretic prediction agents do not always sabotage each other, and yield a Pareto-superior outcome. With regard to the contest stage we find that the exertion of effort is qualitatively in line with theory, i.e. effort is increasing in the sum of marginal costs of an agent's competitors and decreasing in one's own marginal cost of effort. Besides, effort is much higher than predicted by theory. We find indicators that the way of emergence of marginal costs influences the behavior of agents. It seems as if an intentionally well-directed sabotage activity discourages agents more than an ex ante discrimination via the heterogeneity parameter.

In an analysis of three additional treatments it becomes apparent that the revelation of the saboteur's identity has a major influence on agents' behavior. Agents tend to retaliate sabotage activities in the *Info* settings. As conjectured we find that sabotage is higher in asymmetric contests if the identity of saboteurs is not revealed. This is due to the two-person group of agents of the same type in each asymmetric treatment whose sabotage activities are higher without revelation of saboteurs.

We investigate several behavioral hypotheses regarding the sabotage activities in the setting without information on the saboteur's identity as well as in the *Info* treatments. We can confirm each of our behavioral hypotheses on sabotage in at least one setting. Thus, we find that not only the own type of player is decisive for the behavior of agents but the composition of different types of competitors in a contest. The revelation of the saboteur's identity may systematically strengthen respectively weaken these effects.

Literature

Abbink, K. and A. Sadrieh (1995) RatImage – Research Assistance Toolbox for Computer-Aided Human Behavior Experiments, SFB Discussion Paper B-325, University of Bonn.

Baker, G.P., Gibbs, M. and B. Holmström (1994): The Wage Policy of a Firm, Quarterly Journal of Economics 109, 921-955.

Baye, M. R. and H. Hoppe (2003): The Strategic Equivalence of Rent-Seeking, Innovation, and Patent-Race Games, Games and Economic Behavior 44, 217-226.

Becker, B.E. and M. A. Huselid (1992): The Incentive Effects of Tournament Compensation Systems, Administrative Science Quarterly 37, 336-350.

Bull, C., Schotter, A. and K. Weigelt (1987): Tournaments and Piece Rates: An Experimental Study. Journal of Political Economy 95, 1-3.

Chen, K.-P. (2003): Sabotage in Promotion Tournaments. Journal of Law, Economics, and Organization 19 (1), 119-139.

Ehrenberg, R.G. and M. L. Bognanno (1990a): The Incentive Effects of Tournaments Revisited: Evidence from the European PGA Tour, Industrial and Labor Relations Review, Special Issue 43, 74-88.

Ehrenberg, R.G. and M. L. Bognanno (1990b): Do Tournaments Have Incentive Effects?, Journal of Political Economy 98, 1307-1324.

Davis, D. and R. J. Reilly (1998): Do Too Many Cooks Always Spoil the Stew? An Experimental Analysis of Rent-seeking and the Role of a Strategic Buyer, Public Choice 95, 89-115.

Dixit, A. (1987): Strategic Behavior in Contests, American Economic Review 77 (5), 891-898.

Drago, R. W. and G. T. Garvey (1998): Incentives for Helping on the Job. Theory and Evidence, Journal of Labor Economics 16(1), 1-15.

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

Garicano, L. and I. Palacios-Huerta (2000) An Empirical Examination of Multidimensional Effort in Tournaments, Discussion Paper University of Chicago.

Gibbons, R. and K. L. Murphy (1990): Relative Performance Evaluation for Chief Executive Officers, Industrial and Labor Relations Review, Special Issue 43, 30-51.

Gradstein, M. and K. Konrad (1999): Orchestrating Rent-Seeking Contests, Economic Journal 109 (458), 536-545.

Harbring and Irlenbusch (2003a): Incentives in Tournaments with Endogenous Prize Selection, Discussion Paper University of Bonn and University of Erfurt.

Harbring and Irlenbusch (2003b): On the Influence of Wage Compression on Sabotage, Discussion Paper University of Bonn and University of Erfurt.

Harbring and Ruchala (2003): 'Separating the Wheat from the Chaff' – An Experimental Study on the Sorting Function of Tournaments, Working Paper University of Bonn and University of Erfurt.

Konrad, K. (2000): Sabotage in Rent-seeking contests, Journal of Law, Economics and Organization 16(1), 155-165.

Kräkel, M. (2002): U-Type versus J-Type Tournaments, Journal of Institutional and Theoretical Economics 158, 614-637.

Kräkel, M. (2004): Helping and Sabotaging in Tournaments, International Game Theory Review (forthcoming).

Lazear, E. P. (1989): Pay Equality and Industrial Politics. Journal of Political Economy 97, 561-80.

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

Loury, G. C. (1979): Market Structure and Innovation, Quarterly Journal of Economics 93, 395-410.

Lynch, J. and J. S. Zax (1998): Prizes, Selection and Performance in Arabian Horse Racing, Discussion Paper University of Colorado.

Mantrala, M.K., Krafft, M. and B. Weitz (2000): An Empirical Examination of Economic Rationales for Companies' Use of Sales Contests, GEABA Discussion Paper # DP 00-07, German Economic Association of Business Administration.

McLaughlin, K. J. (1988): Aspects of Tournament Models, Research in Labor Economics 9, 225-256.

Millner, E. L. and M. D. Pratt (1991): Risk Aversion and Rent-seeking: An Extension and Some Experimental Evidence, Public Choice 69, 81-92.

Moldovanu, B. and A. Sela, (2001): The Optimal Allocation of Prizes in Contests, American Economic Review, 91, 542-558.

Müller, W. and A. Schotter (2003): Workaholics and Drop outs in Optimal Organizations. Working Paper, New York University.

Potters, J., de Vries, C. and F. van Winden (1998): An Experimental Examination of Rational Rent-Seeking, European Journal of Political Economy 14, 783-800.

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

Szymanski, S. (2004): The Economic Design of Sporting Contests, Journal of Economic Literature 41, 1137-1187.

Tullock, G. (1980): Efficient Rent-seeking, J.M. Buchanan, R.D. Tollison and G. Tullock (Eds.), Toward a Theory of the Rent-seeking Society, 97-112, Texas.

Weigelt, Keith; Dukerich, J. and A. Schotter (1989): Reactions to Discrimination in an Incentive Pay Compensation Scheme: A Game-Theoretic Approach, *Organizational Behavior and Human Decision Processes 44*, 26-44.

Weimann, J., Vogt, C. and C.-L. Yang (2002): Efficient Rent Seeking in Experiments, Public Choice 110, 67-78.

Appendix – Instructions (Example: *HetUInfo*)

Rounds, Groups and Roles

- The experiment consists of 30 rounds. Each round has two stages.
- You will be assigned to a group of three participants. During the whole experiment you will only interact with the participants of your group. The group is randomly composed and stays as such throughout the whole experiment. You will not be informed about who is a participant of your group.
- In every group there are participants of the following colors: Yellow, Red and Blue.
- The participant Blue has lower costs at the second stage than the participants Red and Yellow.
- Costs and payments are stated in the fictitious currency „Taler". At the end of the experiment the sum of every round's payments will be exchanged at a rate of 1 Euro per 1,200 Taler..

Stage 1

- You choose an A or a B for each other participant in your group.
- The number of As and Bs chosen by each participant for both other participants determines the cost level for this round, i.e. each participant's decision at stage 1 influences the cost levels of both other player. In the following all possible combinations of A and B that can be chosen for a participant and the resulting cost levels are presented:

two As	**Cost Level AA**
one A and one B:	**Cost Level AB**
two Bs:	**Cost Level BB**

- After you have chosen an A or a B for both other participants, you are informed about your cost level and the cost level of both other players in your group. Moreover, you are informed about the decisions of the both other players in your groups that led to your cost level.
- The cost for your choice of one A amounts to 0 Taler, and the cost for your choice of one B is 5 Talers.

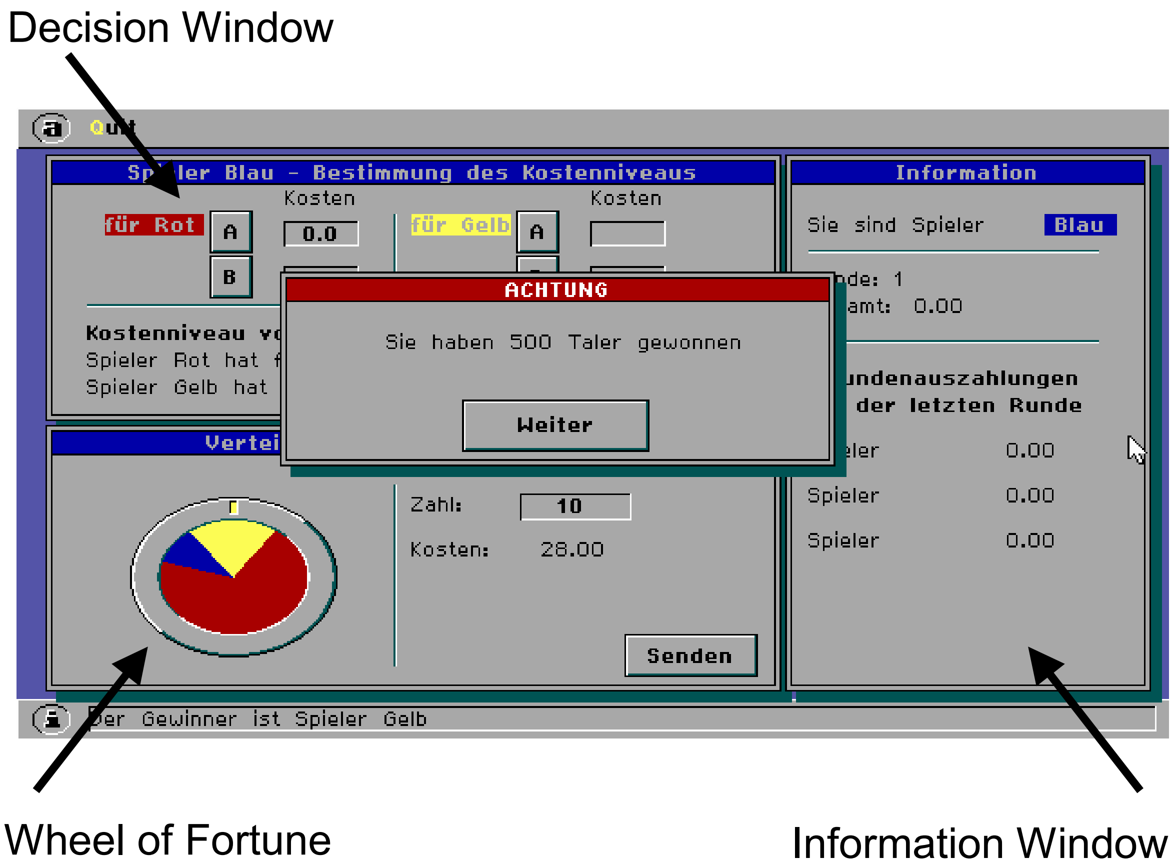
Stage 2

- At stage 2 you choose a number from {0, ..., 300}, which costs a certain amount of Talers.
- Each participant's cost level determines the amount of costs for the number chosen in this round. Costs are lowest on cost level AA, higher on cost level AB and highest on cost level BB. Please learn the cost for the number to be chosen from the corresponding cost table.
- Please note: The participant Blue has lower cost than the participants Red and Yellow. For a specific cost level the costs of a specific number are always lower for participant Blue than for participants Red and Yellow. Attention: The participant Blue thus has different cost tables.
- After all three participants in a group have chosen their numbers, one participant obtains a high payment of 1,500 Talers, and the other two participants receive a low payment of 500 Talers each.
- Your own number and the numbers of both other players determine the probability with which each participant obtains the high payment of 1,500 Talers. The higher the number you choose, the higher is your probability to obtain the 1,500 Talers.

Probability to obtain high payment = Own number / Sum of all three numbers

- After each participant has chosen her number, a „wheel of fortune" determines the participant who obtains the high payment and the two participants who receive the low payments. The colored areas of the wheel of fortune correspond to the probabilities of the participants which result from the chosen numbers.
- Costs of your decision at stage 1 are deducted from your payment as well as costs for the number you have chosen at stage 2. The current round's payment results.

Screenshot (Example)



Cost Table (Example: Underdog at Cost level 2)

Number	Cost	Number	Cost	Number	Cost	Number	Cost	Number	Cost
1	**3.92**	61	**239.12**	121	**474.32**	181	**709.52**	241	**944.72**
2	**7.84**	62	**243.04**	122	**478.24**	182	**713.44**	242	**948.64**
3	**11.76**	63	**246.96**	123	**482.16**	183	**717.36**	243	**952.56**
4	**15.68**	64	**250.88**	124	**486.08**	184	**721.28**	244	**956.48**
5	**19.60**	65	**254.80**	125	**490.00**	185	**725.20**	245	**960.40**
6	**23.52**	66	**258.72**	126	**493.92**	186	**729.12**	246	**964.32**
7	**27.44**	67	**262.64**	127	**497.84**	187	**733.04**	247	**968.24**
8	**31.36**	68	**266.56**	128	**501.76**	188	**736.96**	248	**972.16**
9	**35.28**	69	**270.48**	129	**505.68**	189	**740.88**	249	**976.08**
10	**39.20**	70	**274.40**	130	**509.60**	190	**744.80**	250	**980.00**
11	**43.12**	71	**278.32**	131	**513.52**	191	**748.72**	251	**983.92**
12	**47.04**	72	**282.24**	132	**517.44**	192	**752.64**	252	**987.84**
13	**50.96**	73	**286.16**	133	**521.36**	193	**756.56**	253	**991.76**
14	**54.88**	74	**290.08**	134	**525.28**	194	**760.48**	254	**995.68**
15	**58.80**	75	**294.00**	135	**529.20**	195	**764.40**	255	**999.60**
16	**62.72**	76	**297.92**	136	**533.12**	196	**768.32**	256	**1003.52**
17	**66.64**	77	**301.84**	137	**537.04**	197	**772.24**	257	**1007.44**
18	**70.56**	78	**305.76**	138	**540.96**	198	**776.16**	258	**1011.36**
19	**74.48**	79	**309.68**	139	**544.88**	199	**780.08**	259	**1015.28**
20	**78.40**	80	**313.60**	140	**548.80**	200	**784.00**	260	**1019.20**
21	**82.32**	81	**317.52**	141	**552.72**	201	**787.92**	261	**1023.12**
22	**86.24**	82	**321.44**	142	**556.64**	202	**791.84**	262	**1027.04**
23	**90.16**	83	**325.36**	143	**560.56**	203	**795.76**	263	**1030.96**
24	**94.08**	84	**329.28**	144	**564.48**	204	**799.68**	264	**1034.88**
25	**98.00**	85	**333.20**	145	**568.40**	205	**803.60**	265	**1038.80**
26	**101.92**	86	**337.12**	146	**572.32**	206	**807.52**	266	**1042.72**
27	**105.84**	87	**341.04**	147	**576.24**	207	**811.44**	267	**1046.64**
28	**109.76**	88	**344.96**	148	**580.16**	208	**815.36**	268	**1050.56**
29	**113.68**	89	**348.88**	149	**584.08**	209	**819.28**	269	**1054.48**
30	**117.60**	90	**352.80**	150	**588.00**	210	**823.20**	270	**1058.40**
31	**121.52**	91	**356.72**	151	**591.92**	211	**827.12**	271	**1062.32**
32	**125.44**	92	**360.64**	152	**595.84**	212	**831.04**	272	**1066.24**
33	**129.36**	93	**364.56**	153	**599.76**	213	**834.96**	273	**1070.16**
34	**133.28**	94	**368.48**	154	**603.68**	214	**838.88**	274	**1074.08**
35	**137.20**	95	**372.40**	155	**607.60**	215	**842.80**	275	**1078.00**
36	**141.12**	96	**376.32**	156	**611.52**	216	**846.72**	276	**1081.92**
37	**145.04**	97	**380.24**	157	**615.44**	217	**850.64**	277	**1085.84**
38	**148.96**	98	**384.16**	158	**619.36**	218	**854.56**	278	**1089.76**
39	**152.88**	99	**388.08**	159	**623.28**	219	**858.48**	279	**1093.68**
40	**156.80**	100	**392.00**	160	**627.20**	220	**862.40**	280	**1097.60**
41	**160.72**	101	**395.92**	161	**631.12**	221	**866.32**	281	**1101.52**
42	**164.64**	102	**399.84**	162	**635.04**	222	**870.24**	282	**1105.44**
43	**168.56**	103	**403.76**	163	**638.96**	223	**874.16**	283	**1109.36**
44	**172.48**	104	**407.68**	164	**642.88**	224	**878.08**	284	**1113.28**
45	**176.40**	105	**411.60**	165	**646.80**	225	**882.00**	285	**1117.20**
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.