

Dato, Simon; Nieken, Petra

Conference Paper

Gender Differences in Competition and Sabotage

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2013: Wettbewerbspolitik und Regulierung in einer globalen Wirtschaftsordnung - Session: Experiments on Gender Differences, No. G05-V2

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Dato, Simon; Nieken, Petra (2013) : Gender Differences in Competition and Sabotage, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2013: Wettbewerbspolitik und Regulierung in einer globalen Wirtschaftsordnung - Session: Experiments on Gender Differences, No. G05-V2, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/79750>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Gender Differences in Competition and Sabotage*

Simon Dato[†]

Petra Nieken[‡]

Abstract

We study the differences in behavior of males and females in a two-player tournament with sabotage in a controlled lab experiment. Implementing a real-effort design and a principal who is paid based on the agent's output, we find that males and females do not differ in their achievements in the real effort task but in their choice of sabotage. Males select significantly more sabotage leading to an on average higher winning probability but not to higher profits. If the gender of the opponent is revealed before the tournament, males increase their performance in the real effort task especially if the opponent is female. The gender gap in sabotage is persistent. We discuss possible explanations for our findings and their implications.

Keywords: gender, sabotage, tournament, real effort task
JEL Codes: C91, J16, M52

*We would like to thank the participants of the research seminar of the University of Tübingen, the participants of the Colloquium of Personnel Economics 2013, and in particular Matthias Kräkel and Daniel Müller for helpful comments as well as Torsten Held for programming the experimental software. Financial support by the Deutsche Forschungsgemeinschaft (DFG), grant SFB/TR 15, is gratefully acknowledged.

[†]University of Bonn, Department of Economics, Adenauerallee 24-42, D-53113 Bonn, Germany, tel: +49 228 739214, fax: +49 228 739210, e-mail: simdato@uni-bonn.de.

[‡]University of Bonn, Department of Economics, Adenauerallee 24-42, D-53113 Bonn, Germany, tel: +49 228 739213, fax: +49 228 739210, e-mail: petra.nieken@uni-bonn.de.

1 Introduction

Although making up nearly half of the workforce, it is a well known fact that women are underrepresented in upper hierarchy levels in companies worldwide. In 2009, less than three percent of the largest companies in the European Union were run by women and the share of female board members in these companies was lower than eleven percent. A similar pattern can be observed in the United States: Only twelve CEOs from Fortune500 companies in 2011 were female, resulting in a share of 2.4 percent¹. Researchers have offered several explanations for this observation such as labor market discrimination, differences in education, or preferences as well as the reluctance of females to enter competitions (such as promotion tournaments). Many studies have shown that male participants react stronger to competitive incentives (e.g. Gneezy et al. (2003)) while women have a tendency to abstain from competition and prefer wage schemes with absolute instead of relative compensation (see Croson and Gneezy (2009) for an overview). Niederle and Vesterlund (2007) show that, although having the same ability as men, women shy away from competition. While this finding is partly driven by differences in the participants beliefs about their relative ability there still is a large unexplained gap between males and females. The authors therefore conclude that a general gender difference in preferences towards competition exists. In their review paper Croson and Gneezy (2009) discuss possible explanations why women are less competitive than men and report findings of several experiments with conflicting results. While some studies provide evidence that the differences in behavior can be attributed to social conventions and "nature" (see e.g. Harbaugh et al. (2003), or Gneezy et al. (forthcoming)) others have argued that the differences are due to biological distinctions (see e.g. Bateup et al. (2002), Buser (2012), or Colarelli et al. (2006)). Croson and Gneezy (2009) conclude on page 467 "that both nature and nurture are responsible for the gender differences in competition. [...] Further research is clearly needed."

Our paper contributes to the literature dealing with gender differences in competition by studying the actions of contestants in tournaments if manipulation of output is possible. While it was not possible to manipulate the outcome of the tournament in the afore mentioned studies, this type of unethical behavior is not rare in firms and especially relative performance evaluation schemes are quite prone to such activities. If females are reluctant to manipulate the output of their opponent but anticipate to suffer from a manipulation inflicted on them, this might be a possible motive to avoid competitive situations or for instance not to apply for a suitable job. We conducted a real effort experiment where two players participated in a rank order tournament and had the opportunity to sabotage each other and destroy a certain amount of output of their opponent. To come closer to real world situations, we also had principals who were paid based on the output of the two contestants. Hence, sabotaging not only affected the opponent but also reduced the payment of the principal. We conducted four different treatments: the baseline, the belief, the cheating, and the gender treatment. In the belief treatment we elicited the beliefs about the performance as well as the chosen sabotage of the respective opponent. Similar to Niederle and Vesterlund (2007) we can therefore analyze if those beliefs differ between males and females and drive their behavior. The cheating treatment allows us to check if social

¹See <http://money.cnn.com/magazines/fortune/fortune500/2011/womenceos/>

preferences regarding the principal affected behavior because in contrast to the baseline setting the contestants could increase their own score instead of sabotaging their opponent. This of course enhanced the income of the principal as well and created a positive externality. In the gender treatment we revealed the gender of the opponent before the tournament to study if the behavior changed depending on the gender of the respective opponent. We find that (i) females and males performed equally well in the real effort task if the gender of the opponent was not revealed but (ii) males chose significantly higher levels of sabotage than women. Males, therefore were much more likely to win, especially when competing against a female. Despite this difference, males and females received similar payments at the end as sabotage was costly. In addition, the gender gap in sabotage can also be found in the stated beliefs about the actions of the opponent. If the gender of the opponent was revealed, males increased performance especially if their opponent was female but we do not find any differences in sabotaging behavior with respect to the revealed gender of the opponent. However, the gender gap in sabotaging was persistent over all treatments and women expected more sabotage from males than from females. We discuss several possible explanations for our findings such as the influence of risk attitudes, human values, social preferences, or joy of winning effects.

2 Related Literature

Our paper is related to the literature on sabotage in tournaments. In his seminal paper Lazear (1989) shows that the optimal wage spread is lower when participants are able to sabotage each other. Hence, the tournament designer optimally uses a more equitable prize structure in order to mitigate the incentives to sabotage the opponent.² Since company data on sabotage is basically not available for research, the studies use sports data or experimental data to investigate the impact of sabotage on tournaments. Garicano and Palacios-Huerta (forthcoming) use data from professional soccer and show that a higher prize spread leads to more sabotage activities. Deutscher et al. (forthcoming) report that underdogs exert more sabotage than favorites using a dataset from German professional soccer. Regarding the experimental literature, Harbring and Irlenbusch have contributed several papers dealing with different prize spreads as well as a varying number of participants and different numbers of tournament prizes (Harbring and Irlenbusch (2008)), and communication regarding tournaments with the possibility to sabotage (Harbring and Irlenbusch (2011)). In these experiments, the sabotage activity of a contestant reduces the output of all other contestants. In contrast to this, Harbring et al. (2007) study a heterogeneous tournament with three players, where prior to choosing effort, every player is able to sabotage each opponent separately. The authors show that the observed sabotage activities systematically vary with the composition of the group. The experimental results from Gürtler et al. (forthcoming) indicate that sabotage does not only destroy output but also leads to a reduction of productive effort in sequential tournaments if the contestants expect to be sabotaged. All the above mentioned papers use a chosen effort setting and we are only aware of two papers implementing a real effort tournament with sabotage which are closer to our study. Vandegrift and Yavas (2010) use a forecasting task and gave the contestants the possibility to

²For further theoretical work on sabotage in tournaments see among others Chen (2003), Gürtler (2008), Kräkel (2005), Münster (2007).

raise the forecasting error of their opponent. Sabotage was costly and the authors induced linear costs of sabotage as each unit of raising the error cost a constant fee. They find that players exert more sabotage if the prize spread is higher or the players are rematched each round. In the study of Carpenter et al. (2010) the task was to prepare letters and envelopes. The participants were matched in groups of eight players and each of them counted and evaluated the quality of the work of the other group members. The authors conducted different treatments with piece rate and tournament incentives with and without sabotage. They find that output declines in the tournaments with sabotage compared to the piece rate condition. The reduction is due to false reporting of the quality and not the quantity of the output. Hence, the participants preferred the more subtle form of sabotage and refrain from "undercounting" the total output of an opponent. In contrast to our paper, the participants in the setting of Carpenter et al. (2010) selected their amount of sabotage after the production period where they already knew their own output. Hence, they study a sequential tournament. Furthermore, they had the opportunity to either sabotage the quantity or the quality of the work and Carpenter et al. (2010) argue that both activities led to different psychological costs. However, the authors cannot estimate the cost functions of the participants which are likely to differ between individuals. Our work differs from both papers since we implemented a principal in our setting which induced a negative externality of sabotage on an uninvolved player and we had convex costs of sabotage which were identical for all agents. Furthermore, our study concentrates on the impact of gender on sabotage decisions which is why we conducted different treatments with and without revelation of the opponents' gender as well as positive and negative externalities.

Our work is also related to the growing literature on gender differences in competition. One strand of the literature analyzes the entry decision and studies the question whether women prefer different incentive schemes than men when they are able to choose their compensation scheme. Niederle and Vesterlund (2007) show that women, compared to equally able men, refrain from competition and instead select a piece-rate scheme whereas the majority of male participants enters the tournament. This self-selection effect seems to be very robust, since this finding was replicated by several studies using a similar setup (see e.g. Wozniak et al. (2010), Healy and Pate (2011)) as well as by other studies using different designs (see e.g. Vandegrift and Brown (2005), Dohmen and Falk (2011), Price (2012), Buser (2012), and Garratt et al. (2013)). Another strand of this literature studies the question whether women and men react differently to competitive payment schemes such as rank order tournaments. Gneezy et al. (2003) conduct an experiment in which participants had to carry out a task (solving mazes) being paid according to different compensation schemes. They find no significant performance difference between women and men given a piece-rate, whereas given a competitive compensation scheme, men solved significantly more mazes than women. This finding was replicated by Gneezy and Rustichini (2004) in a field study concerning the competitiveness of ten year old children. The children first had to run 40 meters on their own to measure the speed and afterwards in different-sex pairs. While there was no performance difference in the speed test, boys improved their performance in competition whereas girls became slightly lower. Similar effects can be found in Masclet et al. (2012) who report that men exert higher levels of effort in a competitive compensation scheme than women with similar ability. Although evidence points

into the direction that men react stronger to competitive payment schemes than women, this finding also seems to depend on the kind of task. Günther et al. (2010) report no performance differences given a gender neutral task while they find women to perform better than men in a "female" task. The authors explain their finding with the so called stereotype threat. Further evidence comes from Iriberry and Rey-Biel (2012) who also find that women only perform worse than men if they believe the task is one where men have an advantage. Cotton et al. (2013) develop a formal model to investigate which of the possible explanations fits to the empirical evidence and suggest two reasons for the gender gap in performance. First, the observation that men increase their performance when entering a contest maybe based on changes of their testosterone or adrenaline levels and second men seem to enjoy competing with others and get a positive utility from winning a contest. Buser (2012) also reports that the willingness to compete of females depends in their hormone level, high levels of progesterone reduce the likelihood to enter competition. Furthermore, Fulton et al. (2006) find evidence females in political election contests are more responsive to expected benefits than males when deciding to run for office. This indicates that men face more on the reward while women take the costs and the odds of winning into account. For instance Wieland and Sarin (2012) show that females base their decision whether to compete or not on rationale expectancies while men focus on the economic reward and use competition as a mean to achieve higher self-esteem. The gender composition of the group seems to be of great importance as well: In Gneezy et al. (2003), the gender gap in performance is significantly higher in mixed-sex tournaments than in single-sex tournaments. Furthermore, Datta Gupta et al. (2013) report that men tend to compete less against men than against women which also might be affected by the afore mentioned self-esteem effects.

While all these papers investigate productive behavior in competitive situations we are only aware of one paper which also studies gender effects with respect to unethical behavior in tournaments. Schwierien and Weichselbaumer (2010) allow contestants to cheat to win a tournament. The players had to solve mazes and could either use an "auto solve option" or a "path verifying option" which showed them if the chosen path was wrong to cheat. With a spy software the authors were able to detect the cheating while the players were not aware that their actions would be observed. Cheating was not costly for the players and the authors find that cheating depends on the performance of a contestant. Lower performing players cheated more than better players. Since women on average performed worse than men they were also more likely to cheat. Hence, the observed gender gap in cheating is driven by differences in performance. In contrast to Schwierien and Weichselbaumer (2010), we study sabotage decisions in tournament where sabotaging is costly. In our setting exerting sabotage not only harmed the opponent but also resulted in lower payment for the principal.

For an extensive overview on gender and competition see Niederle and Vesterlund (2011) or Croson and Gneezy (2009) for a more general survey on gender differences in preferences. Furthermore Dechenaux et al. (2012) provide an recent survey on experimental results in contests and tournaments for details on gender see chapter 8.7.

3 Experimental Design

We implemented a two-player real effort tournament with the option to sabotage the opponent. In contrast to chosen effort, real work "involves effort, fatigue, boredom, excitement, and other affections" (Van Dijk et al. (2001), p. 189). Regarding the decision to sabotage and destroy output, we feel that it is important that output is based on real work rather than on a chosen number. In line with Carpenter et al. (2010) and Vandegrift and Yavas (2010) we, therefore, prefer a real effort setting. The experiment encompassed four different treatments named: baseline, belief, cheating, and gender treatment. Each treatment consisted of 8 working periods which lasted 5 minutes each. One of those periods was randomly selected for payment. In each session of a treatment, we had 21 participants who were divided into three units. Each unit contained one principal and six agents. Similar to Erkal et al. (2011) the participants task was to code words into numbers. They were asked to enter a two-digit number for each letter of a word according to an encryption table that assigned a number to each letter of the alphabet. Note that all players received the same words in the same order. We start our description with the baseline treatment each agent was matched with a different agent from his unit at the beginning of every period and had to decide about the amount of sabotage he wanted to inflict on his opponent. We implemented this stranger matching to prevent reputation effects and reciprocity among the agents. He could select an integer $x \in [0, 70]$ which would be deducted from the other contestants number of achieved points. Sabotaging the opponent was costly according to $c(x) = x^2/14$. After both contestants had decided on the amount of sabotage, the five minute working period started. Agents received one point for each correctly coded letter and could proceed to coding the next word only if all letters had been coded correctly. The points earned by coding were summed up after each period. Hence, an agent's output was given by the the number of achieved points minus the suffered sabotage. At the end of each period, the agent with the higher output received 500 taler while the agent with the lower output received 200 taler. Agents were not informed about the amount of sabotage inflicted on them, they only received information on the number of achieved points in the coding task, their own choice of sabotage, and whether they achieved the higher output or not. The principal's payment was determined by the average output of the six agents in the corresponding unit. The principal observed the output of each agent in her unit but not the amount of sabotage or the points achieved in the coding task. The agents were aware that their output determined the payment of the principal and they knew which information she was given. After the experiment, all participants filled out a questionnaire containing questions regarding the experiment. Additionally, we also collected socioeconomic information, basic human values and elicited the risk attitudes of the participants using a question from the GSOEP (for a discussion that this question is a good predictor for actual behavior please refer to Dohmen et al. (2011b)). We used the Schwarz Human Values Questionnaire as implemented in the European Social Survey with 21 items to measure basic human values. According to Schwartz (2000) those values are recognized across different cultures and Societies. Those values are power, achievement, hedonism, stimulation, self-direction, universalism, benevolence, tradition, conformity, and security. For a detailed discussion about human values and gender differences please see Schwartz and Rubel (2005) and Adams and Funk (2012).

The belief treatment is identical to the baseline treatment, the only difference is that we elicited the agents' belief about the number of achieved points as well as the amount of sabotage of their respective opponent. Each period, agents were asked to estimate their opponents number of achieved points and sabotage decision before choosing their own amount of sabotage. The elicitation of beliefs was incentivized, for every stated belief being Z points away from the correct value, agents received $\max\{15 - Z, 0\}$ taler. Hence, based on this treatment we are able to study of the incentivized elicitation of the beliefs at the beginning of each period altered behavior in the tournament. In the gender treatment, again everything else being equal to the belief treatment, the gender of the opponent was revealed prior to the elicitation of beliefs. In the cheating treatment we switched from a sabotage opportunity to a cheating opportunity. In contrast to the baseline treatment, now the chosen amount of points x were added to the own number of achieved points in the coding task instead of being deducted from the opponents number of achieved of points.³

The experiment was conducted at the BonnEconLab. We used the online recruitment ORSEE (Greiner (2004)) and programmed the experimental software in z-tree (Fischbacher (2007)). Each treatment encompassed 4 sessions which lasted 1.5 hours. 336 students enrolled at the University of Bonn participated in the experiment, 192 female and 144 male.⁴ The average earnings were 17.49 euro (including an endowment of 6 euro for agents and 4 euro for principals).

4 Results and Discussion

In our analysis, we focus on the behavior of the participants in the role of agents since the principals had no decision power and only received the produced output in our setting. We start the analysis by looking at the number of achieved points in the coding task in the baseline-treatment. On average females achieved 111.43 points while males achieved 114.87 points. We do not find a significant difference in performance between male and female agents (Mann-Whitney U tests with data pooled over all periods for each agent as well as tested for each period separately). Hence, regarding performance, our results are in line with the findings of Healy and Pate (2011), Günther et al. (2010), Niederle and Vesterlund (2007), and Wozniak et al. (2010). While we observe similar performance levels of males and females in the coding task, we find an improvement of performance over the course of the experiment for both genders (see Figure 1). This is supported by the regressions in columns (2) and (4) of Table A1 in the appendix⁵, as gender has no significant impact on the number of achieved points but the dummy variables for each period have a significant positive impact indicating learning effects.

Next, we inspect the sabotage decisions of the agents in the baseline treatment. On aver-

³Note that in the baseline and the cheating treatment, at the end of the experiment we asked the agents to report their belief about the average amount of x the other agents had chosen as well as if they belief to have selected a higher, an equal or a lower amount of x than their opponent on average.

⁴The instructions of the baseline treatment translated into english can be found in the appendix.

⁵We apply OLS and Random Effects regressions and calculate robust standard errors clustered on subjects. Additionally, we ran those regressions with robust standard errors clustered on matching group level to control for possible biases due to the matching protocol. Note that such a bias is unlikely since the agents did not receive information on the performance or sabotage levels of their opponents. The results remain qualitatively the same and can be obtained upon request.

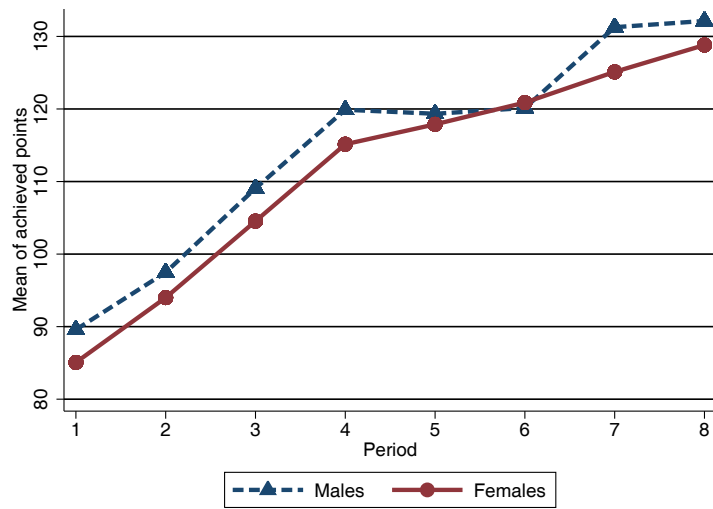


Figure 1: Mean of achieved points in each period for males and females in the baseline treatment

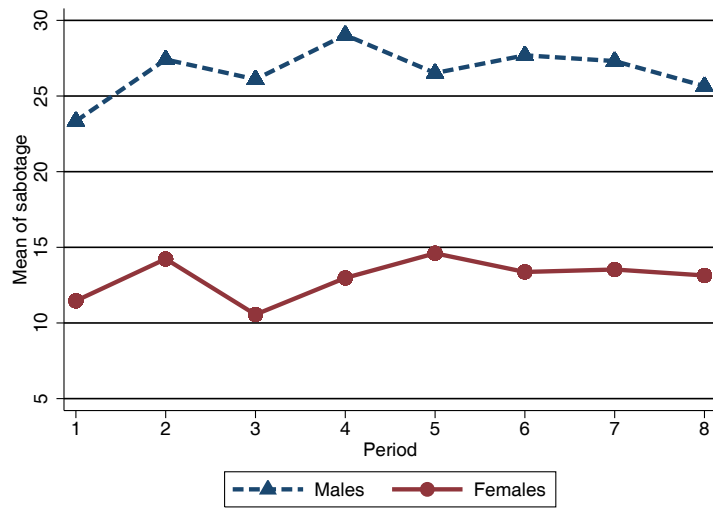


Figure 2: Mean of sabotage in each period for males and females in the baseline treatment

age males selected 26.63 points of sabotage while females selected 12.99 points of sabotage on average. The observation that females selected significantly less sabotage than males is supported by Mann-Whitney U tests using pooled data over all periods ($p = 0.000$) as well as for each period ($p \leq 0.01$ for each period). A closer look reveals that we have 120 observations of females who selected zero sabotage while this can only be observed for 43 male observations. Hence, females more often preferred not to sabotage at all. But even if we only compare the observations where a positive amount of sabotage was chosen, males were more prone to larger amounts of sabotage than females (average sabotage for males 32.69 points and for females 19.94 points). Again we report the results of OLS and Random Effects regressions with robust standard errors clustered on subject level (see Table 1)⁶. Since the amount of sabotage could range between zero and 70, we also apply tobit regressions (see columns (3) and (4) of Table 1). The dummy for females is negative and highly significant in all specifications. As can be seen in Figure 2, there is a slight rise of sabotage during the first periods of the experiment. The regressions in Table 1 reveal that the increase is significant and the gender gap in sabotage is persistent. Hence, while we do not find different levels of performance in the coding task, males selected on average twice as much sabotage than females.

In the following we discuss several possible explanations for this highly significant gender gap in sabotage. We will discuss whether these differences can be explained by risk attitudes, basic human values, or aspects of social preferences. Furthermore, we will analyze the beliefs about the actions of the respective opponent to see if the gender gap was also present in the beliefs. In a last step we are going to report the data of the gender treatment to see if the behavior changed when the gender of the respective opponent was known.

Since previous studies report gender differences regarding the risk attitude (see e.g. Dohmen et al. (2011b) and Holt and Laury (2002)), we checked the impact of the general risk attitude of the agents (measured by the question taken from the GSOEP) on the amount of chosen sabotage but find no significant effect (see columns (2), (4), and (5) of Table 1). We do not find significant differences between males and females regarding risk attitudes in our sample either. This might be due to the fact, that students are a rather homogeneous group regarding other factors such as intelligence or age which are also known to influence risk attitudes besides gender. We have also checked interaction effects between gender and risk attitudes in the regressions but again find no significant impact.

Furthermore, differences in basic human values might have influenced the decision about sabotage. We expect power, achievement, benevolence, and security to be of interest in our setting. Schwartz and Rubel (2005) have shown that males and females differ in the importance of some human values. While power, stimulation, hedonism, achievement, and self-direction seem to be of higher importance for males, benevolence and universalism are more important for females in the general population. Adams and Funk (2012) confirm differences in power, benevolence, and universalism also for male and female board members. However, similar to risk attitudes we find no significant differences regarding basic human values between males and females in the baseline treatment. As can be seen in Table A2 in the appendix the basic human

⁶We also ran those regressions with robust standard errors clustered on matching group level. The results are stable and available upon request.

values also have no significant impact on the chosen sabotage.

	OLS		Tobit		Random Effects
	(1)	(2)	(3)	(4)	(5)
Dummy female	-13.65*** (3.684)	-13.18*** (3.667)	-17.66*** (5.233)	-17.10*** (5.148)	-13.18*** (3.667)
Risk attitude		0.530 (0.756)		0.571 (1.043)	0.530 (0.756)
Dummy period 2		3.292** (1.496)		3.909* (2.138)	3.292** (1.496)
Dummy period 3		0.569 (1.519)		-0.298 (2.319)	0.569 (1.519)
Dummy period 4		3.194** (1.333)		2.864 (1.892)	3.194** (1.333)
Dummy period 5		3.153* (1.734)		2.286 (2.533)	3.153* (1.734)
Dummy period 6		2.889 (1.841)		2.400 (2.625)	2.889 (1.841)
Dummy period 7		2.833* (1.645)		1.702 (2.402)	2.833* (1.645)
Dummy period 8		1.931 (1.643)		0.688 (2.430)	1.931 (1.643)
Constant	26.63*** (3.262)	21.59*** (4.741)	25.14*** (4.177)	20.42*** (6.390)	21.59*** (4.741)
Observations	576	576	576	576	576
# of left censored			163	163	
# right censored			21	21	
Pseudo Log Likelihood			-1963.096	-1961.0809	
(Within) R^2	0.137	0.147			0.0168

Dependent variable is sabotage. Robust standard errors clustered on subjects in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 1: OLS, Tobit, and Random Effects Regressions with sabotage as the dependend variable.

Since in our setting sabotage affected the opponent's expected payoff as well as the principal's payoff, differences in social preferences between females and males could help to explain the gender gap in sabotage. First, sabotage reduced the opponent's output and therefore his chances of winning as well as his expected payoff. Hence, social preferences of the agents with respect to the opponent might be important. As a tournament always leads to an unequal outcome, the behavior of agents could be influenced by inequity aversions (namely envy and

compassion). Grund and Sliwka (2005) have shown that inequity leads to higher levels of sabotage of an inequity averse agent compared to the sabotage level of a self-interested agent if disadvantageous inequity has a higher impact on utility than advantageous inequity. In other words, to explain our findings males would have to be far more prone to envy than females which was not the case in the baseline treatment. However, according to Grund and Sliwka (2005) this should also lead to more productive effort of envious agents and we would expect a higher number of achieved points from males. However, in our setting inequity aversion towards the opponent does not seem to be the driving factor behind our findings.

Second, sabotage affected the payoff of another, uninvolved player, namely the principal: Note that in our setting, similar to many real world situations, sabotage led to a negative externality. Decreasing the output of the opponent not only led to a relative advantage in the tournament but also reduced the payoff of the principal as sabotage destroyed output. If females were less selfish and cared for the principal's payoff to a higher degree than males, they should have chosen lower levels of sabotage. Several experimental results exist, indicating that indeed females are less selfish than males, see for instance Eckel and Grossman (1998), or Güth et al. (2007). Based on these findings, one might believe gender differences in social preferences to be a driving force of the gender gap in sabotage.

To test this hypothesis, we use the results from the cheating treatment. Again, agents chose an integer $x \in [0, 70]$ prior to coding words, but now the chosen amount was added to ones own number of achieved points instead of being deducted from the opponents number of achieved points in the coding task. In this treatment choosing x increased the agents output without exerting higher effort. Note that x affected the agents payoffs in the same way as in the baseline treatment since it also increased the probability to win the tournament and induced the same cost. The only difference was that cheating was output-enhancing rather than being destructive and therefore imposed a positive externality on the principal's payoff. Hence, if differences in social preferences were the driving force of the gender difference in the baseline treatment, we should observe a reverse pattern in the cheating treatment with females choosing a higher amount of x than males.

Again, we find no significant difference regarding the number of achieved points in the coding task between males (on average 107.03 points) and females (on average 111.83 points). Note that the performance is similar to the results in the baseline treatment. The average amount of cheating is 25.38 points for male and 14.86 points for female agents which again is a significant difference according to a Mann-Whitney U test ($p = 0.002$ pooled over all periods, $p \leq 0.03$ for each period except period 2 which is not significant with $p = 0.1543$). As in the baseline treatment, males selected significantly higher amounts of sabotage than females and therefore stronger social preferences of females cannot explain the different amount of sabotage we observe. Moreover, the externality on the principals payoff seems to be irrelevant for the agents choice of x .

Another driving factor might lie in the expectations about the actions of the respective opponent and based on this the perception about the relative performance. If, for instance, females believed that their opponent's performance was lower than their own, there would have been no need for them to invest in costly sabotage to win the tournament. On the other hand

different expectations about the amount of sabotage by the opponent might have driven males to select higher amounts of sabotage. The results of the belief treatment allow us to analyze the assumptions the agents stated about the number of achieved points and the sabotage of the opponent.

First, we compare the number of achieved points as well as the sabotage decisions with the results of the baseline treatment to see if the incentivized elicitation of the beliefs before each period affected the competition. On average females achieved 105.06 points, and males 110.06 points in the coding task. There is no significant difference in performance between male and female agents and the results are similar to those in the baseline treatment. Concerning the sabotage decisions, females chose 10.60 points of sabotage on average while males selected on average 25.56. As in the baseline treatment males chose significantly higher levels of sabotage than females (Mann-Whitney U tests with pooled data $p = 0.001$ and for each period separately $p \leq 0.01$). Since the results are perfectly in line with those of the baseline treatment, we can conclude that the elicitation of beliefs prior to the tournament did not change behavior.

Next we inspect the belief about the opponent’s number of achieved points in the coding task and we do not find a significant gender difference. Females expected their opponents to achieve on average 117.84 points and males expected their opponent to reach 102.31. The relatively high belief of females is driven by one participant who stated an average belief of 518.75. The medians for males and females are identical with 110 points.

Regarding sabotage, we do find a significant gender difference concerning the beliefs about the opponent’s decision: Females expected on average 15.36 points of sabotage to be inflicted on them while males believed that the opponents selected on average 26.13 points of sabotage. The difference between the beliefs is statistically significant (Mann-Whitney U Test $p = 0.006$ pooled over all periods and $p \leq 0.02$ for each period). While we find no gender gap in the performance in the coding task and in the beliefs regarding the performance as well, we find a gender gap in the amount of sabotage as well as in the beliefs about sabotage.

Hence, if there were no significant differences in the performance dimension but males expected a higher amount of sabotage from their opponent, they might have invested higher amounts of sabotage in order to win the tournament. So far, we compared males and females performance and sabotage decisions as well as beliefs about the decisions of the opponent isolated from each other. In a next step, we will calculate the amount of sabotage which was necessary to win the tournament for each agent separately. When deciding on how much to sabotage the opponent, an agent should take into account the belief over the own performance as well as the beliefs about the performance and sabotage of the opponent. Hence, in order to win the sabotage level he chose needed to satisfy

$$\begin{aligned} \text{own sabotage} &\geq \text{belief sabotage opponent} + \text{belief number of points opponent} \\ &\quad - \text{expected number of own points} \end{aligned}$$

Since the agents were informed about their own performance after each period, we use the performance of the previous period as a proxy for the expected own performance. We now calculate the amount \underline{x}_i , which was necessary to at least ensure a tie in the tournament. Based

on the fact that we found no significant difference in the beliefs regarding the performance of the opponent but a significant gender gap in the belief about sabotage, it seems intuitive that \underline{x}_i should be higher for males than for females. However, the \underline{x}_i for males and females was rather similar and we do not find a significant difference according to a Mann-Whitney U test.⁷ This may seem a bit puzzling but can be explained as follows: Although we do not find a statically significant difference between the own (lagged) performance and the belief about the performance of the opponent, males believed to perform slightly better than their opponents. Hence, males seemed to be more optimistic about their relative performance in the coding task than females. On the other hand, they expected more sabotage from their opponents than females. The more optimistic expectation about the relative performance balanced the higher expected amount of sabotage leading to rather similar levels of \underline{x}_i for both genders. However, we still observed higher levels of sabotage from males than from females. Males invested more money to raise their chances to win the tournament. The inspection of the beliefs revealed that the gender gap is also persistent in the beliefs about the sabotage of the opponent. But beliefs alone cannot explain why we find those different beliefs and what drives males to select twice as high levels of sabotage than females.⁸

Cotton et al. (2011) suggest that males derive a higher utility from winning a contest than females because males care more about winning or simply enjoy competitive situations.⁹ They developed a formal model without sabotage which is consistent with previous empirical findings. Wieland and Sarin (2012) show that good performance in a tournament has a stronger impact on the self-worth of males than that of females. They argue that males chose to enter a competition more often in order to enhance the self-esteem and show that winning a tournament is more important to the self-esteem of males than of females. If we transfer these observations to our setting, this can explain why males were willing to invest more money in sabotage even though based on their beliefs one would have expected to see no gender gap in sabotage. Assuming that they derived an additional utility from winning in form of a positive effect on self-esteem or status, it was reasonable for them to select higher amounts of sabotage to raise the odds of winning. If the status of males depends to a higher extent on outperforming opponents and being superior to others (Crocker et al. (2003)), it would be interesting to study how behavior changes if the gender is revealed prior to the competition. Previous findings indicate that males tend to perform slightly better if they compete with a female opponent (Antonovics et al. (2009), Cotton et al. (2013), Price (2008), Iriberry and Rey-Biel (2012)).

The results of the gender treatment allow us to investigate if the agents changed their performance as well as their sabotage when they knew the gender of the opponent. On average males achieved 120.16 points and females 106.05 points in the coding task which is a significant difference in performance (Mann-Whitney U test data pooled over all periods $p = 0.0246$, for each period $p \leq 0.09$, except period 7 with $p = 0.2267$). However, we find no significant

⁷Note that the results do not change if we use the performance of the respective period instead of the previous period as a proxy for the expected own performance.

⁸The interpretation of the beliefs has of course to be regarded with caution since we have point beliefs and do not know how likely the participants thought their stated beliefs were correct. Nevertheless, our data show that we have a gender gap in the beliefs about sabotage as well.

⁹See Dohmen et al. (2011a) for results from neuroscience that are suggestive for the existence of a “joy of winning” effect.

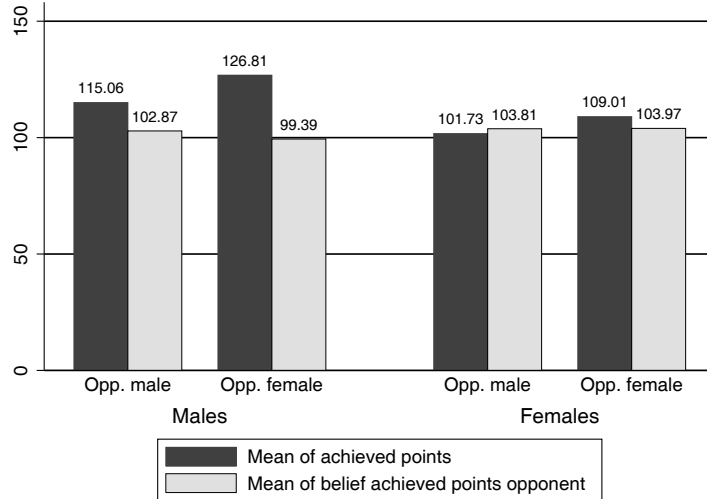


Figure 3: Mean of achieved points in the gender treatment

differences if we compare the performance with the performance in the baseline or the belief treatment. The data indicates that males increased their performance while females did not react so strongly to the revelation of the gender which is in line with previous findings in the literature. Given that we observe a gender gap in performance one might expect that this affected the gender gap in sabotage as well. On average males selected 20.75 points of sabotage and females 10.80 points. Males reduced their sabotage to some extent compared to the other treatments but the gender gap is persistent (Mann-Whitney U Test pooled over all periods $p = 0.0017$, for each period $p \leq 0.04$). Hence, the agents did not change their sabotage decisions. In contrast to all other treatments we now find a significant gender gap in performance. For a more detailed analysis we have to split the sample and investigate the behavior of males and females with respect to the gender of their respective opponent separately. If they anticipated the gender gap in sabotage correctly, we would expect them to adjust their behavior in the tournament accordingly. Figure 3 gives an overview about the performance as well as the beliefs about the performance of males and females when facing a male or female opponent. First, we see that for a given gender of the opponent, males outperformed females. Second, when competing against a female, both genders performed better than when playing against a male opponent which is consistent with previous findings in the literature. The regressions in Table A3 in the appendix support the impression that both the own gender as well as the gender of the opponent influenced the number of achieved points. Both genders increased their performance when competing with a female opponent but the effect is especially prevalent for male agents.¹⁰ We add an interaction effect into the OLS and Random effects regressions but it is only significant in the latter one. The performance difference is even more striking if we look at the beliefs of the male agents. They expected females to perform worse than male

¹⁰If we execute the regressions for males and females separately, the gender of the opponent has a significant effect only in the regressions for males. If we further restrict our sample to agents who have competed with both genders, we have to drop 14 subjects. The results of the regressions are stable with the reduced sample but we cannot observe a significant effect when applying a Wilcoxon matched pair signed rank test. The results can be obtained upon request.

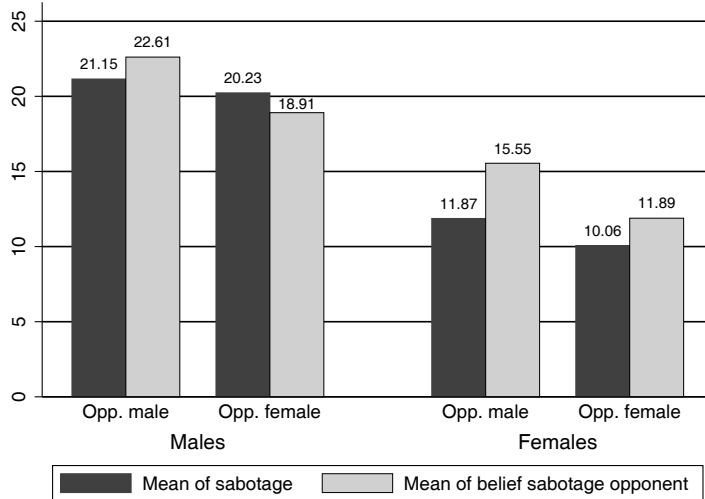


Figure 4: Sabotage in the gender treatment

opponents but nevertheless they increased performance. Losing a competition against a female opponent seems to be worse than losing against a male competitor.

Regarding sabotage, the gender of the opponent has no significant impact on behavior (see Figure 4 and Table A4 in the appendix) and the results are similar as in the baseline and the belief treatment. However, the reported beliefs of the agents differ if they faced a male or female opponent. Especially females expected a higher amount of sabotage from male opponents but this did not lead to higher levels of sabotage if they competed with a male opponent. Even if we do not know why agents expected a higher amount of sabotage from males, this indicates that they were aware that some factors led to differences in the sabotaging behavior of males and females.

The remaining question is whether it paid in monetary terms for males to invest in sabotage and how the gender gap affected the outcome of the tournament. Since both genders on average achieved the same number of points in the coding task (except for the gender treatment) and males chose a higher amount of sabotage, it is straightforward that males won the tournament more often (Two-sided Fisher's exact test $p = 0.000$). Hence, males received the winner prize more frequently but since they had to bear higher costs at the same time the effect on profits is ambiguous. Average earnings in the baseline and belief treatment of males were 18.28 euro and females earned on average 18.22 euro. In the cheating treatment (gender treatment), male agents earned on average 17.77 euro (19.22 euro), whereas female agents received 18.11 euro (17.39 euro).¹¹ We do not find a significant gender difference concerning the earnings in any treatment (Mann-Whitney U test). Hence, the reluctance of females to sabotage their opponents (or to cheat) led to less success in the tournament but did not lead to less profit compared to males. However, if males had an additional utility from winning the tournament, they were better off than females.

From the principals point of view, sabotage reduced his earnings. Furthermore, the relative

¹¹In order to make payoffs comparable between treatments, we adjusted the payoffs from the belief and gender treatment for the incentivized belief elicitation.

performance signal of the agents was biased by sabotage and it remains an open question if principals would expect more sabotage from males and would adjust their expectations about the performance accordingly. If the outcome of the tournament would be used for promotion or sorting decisions, those would systematically favor male agents. Among all male winners in the baseline treatment, about 20.40% won the tournament because of sabotage, whereas this share is 6.62% and accordingly much lower for females.

5 Conclusion

If the gender of the opponent is known, males enhanced their productive effort and achieved more points in the coding task, especially if the respective opponent was female. In all treatments the gender difference in the level of sabotage is large and highly significant. On average males select twice as much sabotage than females. Consequently, males won the tournament more often but since higher amounts of sabotage also led to higher costs, average profit did not differ significantly for both genders. Our data revealed that the gender gap was already present in the beliefs of the players. Social preferences with respect to the principal, risk aversion or human values such as the importance of power or achievement cannot explain the differences in sabotage.

Based on previous findings and the data of the belief- and the gender-treatment, we think it is likely that males derived some extra utility of winning the competition and were therefore willing to invest money to ensure their victory. Females on the other hand are usually not so status seeking. They, therefore, were not willing to invest so much in sabotage. However, they were aware that on average their opponents would choose higher levels of sabotage and that males on average chose higher levels of sabotage than female opponents.

Our results indicate that the revelation of the opponents gender led to a higher performance of males while leaving sabotage almost unchanged. This raised the total output and the behavior was strongest if the opponent was female. Hence, a principal can achieve higher output levels in mixed tournaments where the gender of the opponent is known. On the other hand, principals should of course also try to prevent sabotage as it leads to distorted outcomes. While this recommendation is not new in the literature, we provide an additional reason why sabotage is harmful for organizations and that it can lead to systematically wrong promotion decisions. In our data about 20% of the male winners had achieved their victory based in sabotage and not based on performance.

If companies have established gender quotas or think about affirmative actions programs they need to take into account that sabotage might bias the results of tournaments and that the signals are biased at the expense of females. The reluctance of females to compete or to apply for jobs they are well qualified for might also be affected by the fear of being sabotaged. Sabotage reduces the chances to win and therefore it might be a rationale decision to refuse to enter the tournament.

6 Appendix

	OLS			Random Effects	
	(1)	(2)	(3)	(4)	(5)
Dummy Female	-3.436 (7.175)	-3.436 (7.220)	-4.501 (7.285)	-3.436 (7.220)	-4.501 (7.285)
Risk Attitude			-1.194 (1.347)		-1.194 (1.347)
Dummy Period 2		8.514*** (1.916)	8.514*** (1.918)	8.514*** (1.916)	8.514*** (1.918)
Dummy Period 3		19.47*** (1.776)	19.47*** (1.777)	19.47*** (1.776)	19.47*** (1.777)
Dummy Period 4		30.17*** (2.027)	30.17*** (2.029)	30.17*** (2.027)	30.17*** (2.029)
Dummy Period 5		31.58*** (2.258)	31.58*** (2.260)	31.58*** (2.258)	31.58*** (2.260)
Dummy Period 6		33.69*** (2.520)	33.69*** (2.523)	33.69*** (2.520)	33.69*** (2.523)
Dummy Period 7		40.71*** (2.881)	40.71*** (2.884)	40.71*** (2.881)	40.71*** (2.884)
Dummy Period 8		43.26*** (2.584)	43.26*** (2.587)	43.26*** (2.584)	43.26*** (2.587)
Constant	114.9*** (5.706)	88.94*** (5.323)	95.28*** (8.979)	88.94*** (5.323)	95.28*** (8.979)
Observations	576	576	576	576	576
(Within) R^2	0.002	0.174	0.181	0.5938	0.5938

Dependent variable is number of achieved points. Robust standard errors clustered on subjects in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A1: OLS and Random Effects Regressions for the baseline treatment with number of achieved points as the dependent variable.

	OLS	Tobit	Random Effects
Dummy female	-13.61*** (4.270)	-17.91*** (5.749)	-13.61*** (4.270)
Power	0.678 (2.920)	1.464 (3.912)	0.678 (2.920)
Achievement	1.050 (2.145)	1.620 (3.081)	1.050 (2.145)
Benevolence	-1.516 (3.160)	-1.058 (4.280)	-1.516 (3.160)
Security	-0.601 (1.957)	-1.370 (2.609)	-0.601 (1.957)
Dummy period 2	3.717** (1.764)	3.846 (2.481)	3.717** (1.764)
Dummy period 3	0.933 (1.594)	-0.435 (2.383)	0.933 (1.594)
Dummy period 4	2.883* (1.557)	2.114 (2.220)	2.883* (1.557)
Dummy period 5	3.100* (1.721)	1.951 (2.491)	3.100* (1.721)
Dummy period 6	2.233 (1.953)	1.091 (2.756)	2.233 (1.953)
Dummy period 7	2.417 (1.760)	0.687 (2.549)	2.417 (1.760)
Dummy period 8	2.417 (1.641)	1.145 (2.323)	2.417 (1.641)
Constant	24.64*** (4.067)	23.86*** (5.297)	24.64*** (4.067)
Observations	480	480	480
# of left censored		-1603.5838	
# right censored		144	
Pseudo Log Likelihood		12	
(Within) R^2	0.165		0.0162

Dependent variable is sabotage. Robust standard errors clustered on subjects in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A2: OLS and Tobit, and Random Effects regressions for the baseline treatment with sabotage as the dependent variable.

	OLS		Random Effects	
	(1)	(2)	(3)	(4)
Dummy female	-15.63*** (5.465)	-11.59* (6.251)	-14.80*** (5.511)	-11.31* (5.990)
Dummy opponent female	9.539*** (3.227)	13.43*** (4.435)	4.383** (1.774)	7.223*** (2.772)
Female*Opp. female		-7.853 (6.481)		-6.651** (3.185)
Dummy period 2	8*** (1.320)	8.218*** (1.398)	8*** (1.283)	8.185*** (1.331)
Dummy period 3	17.53*** (1.642)	18.18*** (1.787)	17.53*** (1.599)	18.08*** (1.608)
Dummy period 4	28.61*** (1.924)	29.16*** (2.037)	28.61*** (1.780)	29.07*** (1.812)
Dummy period 5	33.12*** (2.534)	33.34*** (2.495)	33.13*** (2.539)	33.31*** (2.479)
Dummy period 6	37.76*** (2.448)	38.20*** (2.541)	37.76*** (2.363)	38.13*** (2.392)
Dummy period 7	44*** (3.033)	44.65*** (3.006)	44*** (3.001)	44.55*** (2.884)
Dummy period 8	43.69*** (3.073)	43.91*** (3.013)	43.69*** (3.078)	43.88*** (3.009)
Constant	89.43*** (4.167)	87.37*** (4.587)	91.66*** (4.002)	90.12*** (4.257)
Observations	576	576	576	576
(Within) R^2	0.309	0.312	0.6018	0.6059

Dependent variable is number of achieved points. Robust standard errors clustered on subjects in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A3: OLS and Random Effects Regressions with number of achieved points in the gender treatment as the dependend variable.

	OLS		Tobit		Random Effects	
	(1)	(2)	(3)	(4)	(5)	(6)
Dummy female	-9.741*** (3.093)	-9.115** (3.785)	-15.25*** (4.612)	-13.91** (5.692)	-9.926*** (3.068)	-8.101** (3.242)
Dummy opponent female	-1.363 (1.795)	-0.761 (2.549)	-1.333 (2.772)	-0.136 (3.207)	-0.202 (1.254)	1.295 (1.804)
Female*Opp. female		-1.215 (3.673)		-2.585 (5.786)		-3.476 (2.429)
Dummy period 2	2.819** (1.387)	2.853** (1.399)	3.040 (2.111)	3.093 (2.106)	2.819** (1.381)	2.916** (1.365)
Dummy period 3	2.875* (1.495)	2.976* (1.563)	2.673 (2.220)	2.895 (2.282)	2.875* (1.486)	3.165** (1.473)
Dummy period 4	4.069** (1.746)	4.154** (1.774)	4.267* (2.510)	4.435* (2.512)	4.069** (1.738)	4.311** (1.687)
Dummy period 5	4.542** (1.763)	4.575** (1.763)	4.696* (2.449)	4.753* (2.428)	4.542** (1.767)	4.638*** (1.742)
Dummy period 6	3.014* (1.725)	3.081* (1.749)	3.006 (2.472)	3.157 (2.471)	3.014* (1.713)	3.207* (1.690)
Dummy period 7	3.458* (1.912)	3.560* (1.965)	3.565 (2.662)	3.761 (2.689)	3.458* (1.906)	3.748** (1.881)
Dummy period 8	3.083 (2.073)	3.117 (2.069)	1.924 (2.939)	1.996 (2.914)	3.083 (2.066)	3.180 (2.035)
Constant	18.36*** (2.505)	18.04*** (2.738)	16.05*** (3.406)	15.42*** (3.600)	17.86*** (2.509)	17.05*** (2.567)
Observations	576	576	576	576	576	576
# of left censored			116	116		
# right censored			3	3		
Pseudo Log Likelihood			-1864.2963	1864.0796		
(Within) R^2	0.100	0.100			0.0200	0.0263

Dependent variable is sabotage. Robust standard errors clustered on subjects in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A4: OLS, Tobit, and Random Effects Regressions with sabotage in the gender treatment as the dependent variable.

Instructions for the baseline treatment:

Welcome to this experiment!

You are participating in an economic experiment. All decisions are made privately, meaning that none of the other participants learns the identity of someone having made a decision. The payment is private information as well; none of the participants learns how much others have earned. Please read the instructions carefully. If you have trouble understanding the instructions, please take a second look at it. If you still have questions, please give us a signal.

Overview

- The experiment consists of 8 identical rounds. At the end, one round will be drawn randomly for payment. After the last round, you will receive an overview of your possible payments.
- In this experiment, you are randomly assigned to a unit with seven members each. Throughout the experiment, you will only play with members from your unit. You and the other participants never learn the identity of the other members.
- There are two types of players: type I and type II. There are six players of type I and one player of type II in each unit.
- You will learn your type at the beginning of the experiment. Your type will not change throughout the whole experiment.
- In each round, every participant encodes words to numbers. You have to replace each letter of a word with the numbers given in table 1. You will earn one point for each letter.
- Before the actual 8 rounds of the experiment begin, you have the opportunity to become familiar with the task in a one minute practice round. The practice round only serves for better understanding of the task and does not affect your payment.
- Within a unit, two players of type I are assigned to one group. Please note that the group members change every round and that the identity of the player remains unknown.
- Within a group, the overall score of both type I players will be compared at the end of each round. The player with the higher overall score earns 500 tokens, the other player earns 200 tokens.
- Players of type II do not make any decisions that affect their own payment or the payment of other players in this experiment.
- At the end of the experiment, you will complete a short questionnaire. When all participants have completed this questionnaire, we will start paying out the participants one after another.

Course of action

- Your task is to encode words to numbers. You have to replace each letter of a word with the numbers given in table 1:
 - Example: The word *hat* appears on the screen. According to table 1, $H = 15$, $A = 21$ and $T = 91$, so the code for *hat* is: 152191.
 - For each letter, you have to enter the assigned number in a separate box. You can go from box to box using the tab key.
- Five-digit, six-digit and seven-digit words will appear. You will earn one point for each letter. Please note that you will only earn points, if you encode the whole word correctly. The sum of the points is the obtained score.
- When you have entered the code and pressed *OK*, you receive a notification whether the word was encoded correctly. If so, please click on *next* in order to receive a new word. If the word was not encoded correctly, please try again until you succeed. You have five minutes working time per round. Thereafter, working time will stop automatically.

Type I:

- In each round, you will be randomly assigned to some other type I player from your unit in a group.
- The other player also encodes words to numbers. Note that you will both receive the same words in the same order.
- At the end of each round, the overall score of both type I players will be compared. The player with the higher overall score receives the higher payment of 500 tokens. The more words a player encoded correctly, the higher the obtained score will be. Please note that your overall score is only used for comparison with the score of the other player. Only if your overall score is higher than the score of the other player, you receive the higher payment of 500 tokens. It does not matter how many points you outperform the other player. If your overall score is lower than the score of the other player, you receive 200 tokens.
- Before a round starts, you have the opportunity to reduce the overall score of the other type I player by the amount of X . In this way, the other type I player has a disadvantage when the overall scores are compared at the end of the round. The more points you deduct from the other player, the higher are your resulting costs. The costs will be deducted from your payment for this round in any case. An overview of the costs can be found in table 2.
- In the same way, the other type I player decides whether he wants to reduce your score before the round starts.

- The overall score of a type I player consists of his obtained encoding score minus the amount of X the other player has chosen. Please note that the other type I player never learns which amount of X you have chosen, nor do you receive any information on the amount of X the other type I player has chosen.

Overall score = obtained encoding score - amount of X (chosen by the other type I player)

- At the end of each round, we will show you the following information:
 - Your obtained encoding score
 - The amount of X you have chosen
 - Your payment, if this round is drawn.
- Thereafter a new round starts.

Type II:

- Type II players have no influence on the overall score of the type I players. You will also encode words to numbers, but you do not receive a special payment for this. At the end of each round, you receive an overview of the overall score of all type I players from your unit. Your payment depends on the average overall score of all six type I players from your unit. One point equals two tokens.

Payment

- At the end of the experiment, the round that determines your payment is drawn randomly.
- During the whole experiment, the payments are shown in the currency *tokens*, which will be converted at the end. The conversion rate is 25 tokens → 1 Euro.

Type I:

- The overall score of both type I players from the allotted round influence their payment and the payment of type II. The type I player from the group with the higher overall score receives 500 tokens, the type I player with the lower overall score receives 200 tokens.
- The costs for the chosen amount of X in the allotted round will be deducted from the payment of each player. This yields the overall payment at the end of the experiment. In addition, each player receives a fix amount of 150 tokens.

Higher overall score: **500 tokens - costs for the score of X + 150 tokens**

Lower overall score: **200 tokens - costs for the score of X + 150 tokens**

Type II

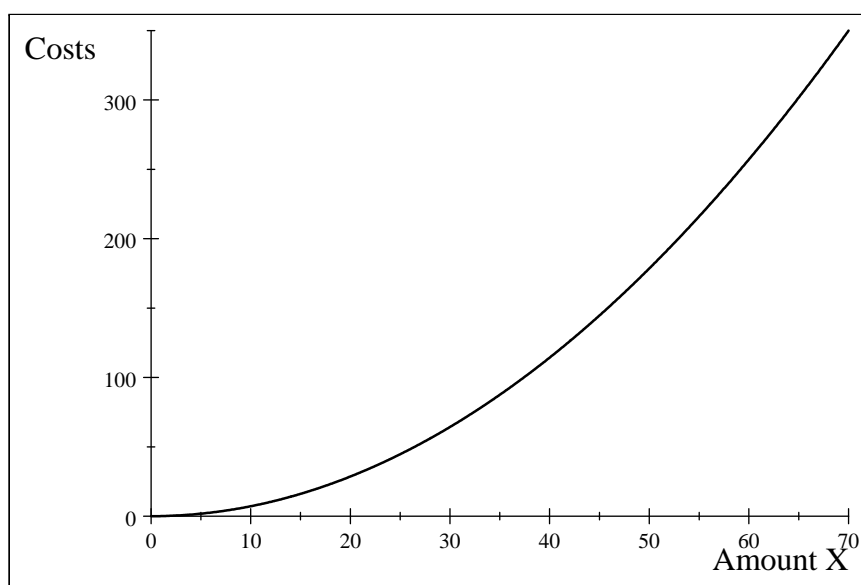
- The type II player receives the average of the obtained overall score of the six type I players from his unit in the allotted round as a payment. One point equals 2 tokens.
2x(average of the unit's overall score) + 100 tokens

Letter	Number	Letter	Number
A	21	N	32
B	54	O	56
C	13	P	10
D	67	Q	23
E	85	R	49
F	31	S	82
G	46	T	91
H	15	U	37
I	98	V	43
J	75	W	52
K	42	X	87
L	27	Y	93
M	19	Z	30

Table A5: Table 1: Overview of the numerical codes to be used

There will be no costs if you choose $X=0$.										
X	1	2	3	4	5	6	7	8	9	10
Costs	0.07	0.29	0.64	1.14	1.79	2.57	3.50	4.57	5.79	7.14
X	11	12	13	14	15	16	17	18	19	20
Costs	8.64	10.29	12.07	14.00	16.07	18.29	20.64	23.14	25.79	28.57
X	21	22	23	24	25	26	27	28	29	30
Costs	31.50	34.57	37.79	41.14	44.64	48.29	52.07	56.00	60.07	64.29
X	31	32	33	34	35	36	37	38	39	40
Costs	68.64	73.14	77.79	82.57	87.50	92.57	97.79	103.14	108.64	114.29
X	41	42	43	44	45	46	47	48	49	50
Costs	120.07	126.00	132.07	138.29	144.64	151.14	157.79	164.57	171.50	178.57
X	51	52	53	54	55	56	57	58	59	60
Costs	185.79	193.14	200.64	208.29	216.07	224.00	232.07	240.29	248.64	257.14
X	61	62	63	64	65	66	67	68	69	70
Costs	265.79	274.57	283.50	292.57	301.79	311.14	320.64	330.29	340.07	350.00

Table A6: Table 2: Overview of the costs for the score X



References

- Adams, R. B. and Funk, P. (2012): Beyond the glass ceiling: Does gender matter? *Management Science*, 58(2), pp. 219–235.
- Antonovics, K., Arcidiacono, P. and Walsh, R. (2009): The Effects of Gender Interactions in the Lab and in the Field. *The Review of Economics and Statistics*, 91(1), pp. 152–162.
- Bateup, H.S., Booth, A., Shirtcliff, E.A. and Granger, D.A. (2002): Testosterone, cortisol, and women’s competition. *Evolution and Human Behavior*, 23(3), pp. 181–192.
- Buser, T. (2012): The impact of the menstrual cycle and hormonal contraceptives on competitiveness. *Journal of Economic Behavior & Organization*, 83(1), pp. 1–10.
- Carpenter, J., Matthews, P. H. and Schirm, J. (2010): Tournaments and Office Politics: Evidence from a Real Effort Experiment. *American Economic Review*, 100(1), pp. 504–17.
- Chen, K. P. (2003): Sabotage in promotion tournaments. *Journal of Law, Economics, and Organization*, 19(1), pp. 119–140.
- Colarelli, S. M., Spranger, J. L. and Hechanova, M. R. (2006): Women, power, and sex composition in small groups: An evolutionary perspective. *Journal of Organizational Behavior*, 27(2), pp. 163–184.
- Cotton, C., McIntyre, F. and Price, J. (2011): Gender Differences in Competition: A Theoretical Assessment of the Evidence.
- Cotton, C., McIntyre, F. and Price, J. (2013): Gender differences in repeated competition: Evidence from school math contests. *Journal of Economic Behavior & Organization*, 86(0), pp. 52–66.
- Crocker, J., Luhtanen, R. K., Cooper, M. L. and Bouvrette, A. (2003): Contingencies of self-worth in college students: theory and measurement. *Journal of personality and social psychology*, 85(5), p. 894.
- Croson, R. and Gneezy, U. (2009): Gender Differences in Preferences. *Journal of Economic Literature*, 47(2), pp. 448–474.
- Datta Gupta, N., Poulsen, A. and Villeval, M. C. (2013): Gender matching and competitiveness: experimental evidence. *Economic Inquiry*, 51(1), pp. 816–835.
- Dechenaux, E., Kovenock, D. and Sheremeta, R. (2012): A survey of experimental research on contests, all-pay auctions and tournaments.
- Deutscher, C., Frick, B., Gürtler, O. and Prinz, J. (forthcoming): Sabotage in Tournaments with Heterogeneous Contestants: A Field Study. *The Scandinavian Journal of Economics*.
- Dohmen, T. and Falk, A. (2011): Performance Pay and Multidimensional Sorting: Productivity, Preferences, and Gender. *American Economic Review*, 101(2), pp. 556–90.

- Dohmen, T., Falk, A., Fliessbach, K., Sunde, U. and Weber, B. (2011a): Relative versus absolute income, joy of winning, and gender: Brain imaging evidence. *Journal of Public Economics*, 95(3), pp. 279 – 285.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J. and Wagner, G. G. (2011b): Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3), pp. 522–550.
- Eckel, C. C. and Grossman, P. J. (1998): Are Women Less Selfish Than Men?: Evidence From Dictator Experiments. *The Economic Journal*, 108(448), pp. 726–735.
- Erkal, N., Gangadharan, L. and Nikiforakis, N. (2011): Relative Earnings and Giving in a Real-Effort Experiment. *American Economic Review*, 101(7), pp. pp. 3330–3348.
- Fischbacher, U. (2007): z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10, pp. 171–178.
- Fulton, S. A., Maestas, C. D., Maisel, L. S. and Stone, W. J. (2006): The Sense of a Woman: Gender, Ambition, and the Decision to run for Congress. *Political Research Quarterly*, 59(2), pp. 235–248.
- Garicano, L. and Palacios-Huerta, I. (forthcoming): Sabotage in Tournaments: Making the Beautiful Game a Bit Less Beautiful. *Beautiful Game Theory*.
- Garratt, R. J., Weinberger, C. and Johnson, N. (2013): The state street mile: Age and gender differences in competition aversion in the field. *Economic Inquiry*, 51(1), pp. 806–815.
- Gneezy, U., Leonard, K. and List, J. (forthcoming): Gender Differences in Competition: The Role of Socialization. *Econometrica*.
- Gneezy, U., Niederle, M. and Rustichini, A. (2003): Performance in Competitive Environments: Gender Differences. *The Quarterly Journal of Economics*, 118(3), pp. pp. 1049–1074.
- Gneezy, U. and Rustichini, A. (2004): Gender and Competition at a Young Age. *The American Economic Review*, 94(2), pp. pp. 377–381.
- Greiner, B. (2004): An Online Recruitment System for Economic Experiments. *Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht*, 63, pp. 79–93.
- Grund, C. and Sliwka, D. (2005): Envy and Compassion in Tournaments. *Journal of Economics and Management Strategy*, 14, pp. 187–207.
- Günther, C., Ekinci, N. A., Schwioren, C. and Strobel, M. (2010): Women can't jump? An experiment on competitive attitudes and stereotype threat. *Journal of Economic Behavior & Organization*, 75(3), pp. 395 – 401.
- Gürtler, O. (2008): On sabotage in collective tournaments. *Journal of Mathematical Economics*, 44(3), pp. 383–393.

- Gürtler, O., Münster, J. and Nieken, P. (forthcoming): Information Policy in Tournaments with Sabotage. *The Scandinavian Journal of Economics*.
- Güth, W., Schmidt, C. and Sutter, M. (2007): Bargaining outside the lab - a newspaper experiment of a three-person ultimatum game. *The Economic Journal*, 117(518), pp. 449–469.
- Harbaugh, W., Krause, K. and Liday, S. (2003): Bargaining by children. University of Oregon Economics Working Paper, (2002-4).
- Harbring, C. and Irlenbusch, B. (2008): How many winners are good to have?: On tournaments with sabotage. *Journal of Economic Behavior & Organization*, 65(3), pp. 682 – 702.
- Harbring, C. and Irlenbusch, B. (2011): Sabotage in tournaments: Evidence from a laboratory experiment. *Management Science*, 57(4), pp. 611–627.
- Harbring, C., Irlenbusch, B., Kräkel, M. and Selten, R. (2007): Sabotage in Corporate Contests - An Experimental Analysis. *International Journal of the Economics of Business*, 14(3), pp. 367–392.
- Healy, A. and Pate, J. (2011): Can Teams Help to Close the Gender Competition Gap? *The Economic Journal*, 121(555), pp. 1192–1204.
- Holt, C. A. and Laury, S. K. (2002): Risk Aversion and Incentive Effects. *The American Economic Review*, 92(5), pp. pp. 1644–1655.
- Iriberri, N. and Rey-Biel, P. (2012): Let’s (not) talk about sex: The effect of information provision on gender differences in performance under competition.
- Kräkel, M. (2005): Helping and sabotaging in tournaments. *International Game Theory Review*, 7(02), pp. 211–228.
- Lazear, Edward P. (1989): Pay Equality and Industrial Politics. *Journal of Political Economy*, 97(3), pp. 561–580.
- Masclet, D., Peterle, E. and Larribeau, S. (2012): Gender Differences in Competitive and Non Competitive Environments: An Experimental Evidence.
- Münster, J. (2007): Selection tournaments, sabotage, and participation. *Journal of Economics & Management Strategy*, 16(4), pp. 943–970.
- Niederle, M. and Vesterlund, L. (2007): Do Women Shy Away From Competition? Do Men Compete Too Much? *The Quarterly Journal of Economics*, 122(3), pp. 1067–1101.
- Niederle, M. and Vesterlund, L. (2011): Gender and Competition. *Annual Review of Economics*, 3(1), pp. 601–630.
- Price, C. R. (2012): Gender, competition, and managerial decisions. *Management Science*, 58(1), pp. 114–122.

- Price, J. (2008): Gender Differences in the Response to Competition. *Industrial and Labor Relations Review*, 61(3), pp. 320–333.
- Schwartz, B. (2000): Self-Determination: The Tyranny of Freedom. *American Psychologist*, 55(1), pp. 79–88.
- Schwartz, S. H. and Rubel, T. (2005): Sex differences in value priorities: cross-cultural and multimethod studies. *Journal of personality and social psychology*, 89(6), p. 1010.
- Schwieren, C. and Weichselbaumer, D. (2010): Does competition enhance performance or cheating? A laboratory experiment. *Journal of Economic Psychology*, 31(3), pp. 241–253.
- Van Dijk, F., Sonnemans, J. and Van Winden, F. (2001): Incentive systems in a real effort experiment. *European Economic Review*, 45(2), pp. 187–214.
- Vandegrift, D. and Brown, P. (2005): Gender differences in the use of high-variance strategies in tournament competition. *The Journal of Socio-Economics*, 34(6), pp. 834 – 849.
- Vandegrift, D. and Yavas, A. (2010): An Experimental Test of Sabotage in Tournaments. *Journal of Institutional and Theoretical Economics JITE*, 166(2), pp. 259–285.
- Wieland, A. and Sarin, R. (2012): Domain specificity of sex differences in competition. *Journal of Economic Behavior & Organization*, 83(1), pp. 151–157.
- Wozniak, D., Harbaugh, W. T. and Mayr, U. (2010): Choices About Competition: Differences by Gender and Hormonal Fluctuations, and the Role of Relative Performance Feedback.