

IZA DP No. 4955

Gender, Competition and the Efficiency of Policy Interventions

Loukas Balafoutas
Matthias Sutter

May 2010

Gender, Competition and the Efficiency of Policy Interventions

Loukas Balafoutas

University of Innsbruck

Matthias Sutter

*University of Innsbruck,
University of Gothenburg and IZA*

Discussion Paper No. 4955

May 2010

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Gender, Competition and the Efficiency of Policy Interventions^{*}

Recent research has shown that women shy away from competition more often than men. We evaluate experimentally three alternative policy interventions to promote women in competitions: *Quotas*, *Preferential Treatment*, and *Repetition of the Competition* unless a critical number of female winners is reached. We find that Quotas and Preferential Treatment encourage women to compete significantly more often than in a control treatment, while efficiency in selecting the best candidates as winners is not worse. The level of cooperation in a post-competition teamwork task is even higher with successful policy interventions. Hence, policy measures promoting women can have a double dividend.

JEL Classification: C91, D03

Keywords: competition, gender gap, experiment, affirmative action, teamwork, coordination

Corresponding author:

Matthias Sutter
Department of Public Finance
University of Innsbruck
Universitaetsstrasse 15
A-6020 Innsbruck
Austria
E-mail: matthias.sutter@uibk.ac.at

^{*} We would like to thank Gary Charness, Nick Feltovich, Uri Gneezy, Martin Kocher, Peter Kuhn and Pedro Rey-Biel as well as seminar participants at the University of Arizona, the University of California at Santa Barbara and the ESA Meetings in Innsbruck and Melbourne for valuable comments. Financial support from the D. Swarovski & Co. Förderfonds of the University of Innsbruck is gratefully acknowledged.

1. Introduction

Despite improvements over the past decades, there are still substantial gender differences on labor markets, such as women lagging behind men with respect to wage levels or opportunities for career advancement (see, for example, Blau and Kahn, 2000; Bertrand and Hallock, 2001; Weichselbaumer and Winter-Ebmer, 2007; Blau, Ferber and Winkler, 2010). These gender differences are often attributed to differences in preferences, problems in combining family and career, or also to discrimination against women (see, e.g., Altonji and Blank, 1999; Goldin and Rouse, 2000; Black and Strahan, 2001). A recent line of research has highlighted another important factor, namely a lower level of competitiveness of women (Gneezy, Niederle and Rustichini, 2003; Gneezy and Rustichini, 2004; Datta Gupta, Poulsen and Villeval, 2005; Niederle and Vesterlund, 2007, 2010; Croson and Gneezy, 2009; Gneezy, Leonard and List, 2009; Wozniak, Harbaugh and Mayr, 2010). These studies provide evidence that men increase their performance under competition more than women¹ and that women more often opt out of competition, even when they are equally qualified. As a consequence of these gender differences in competitiveness, women may get fewer promotion opportunities and subsequently lower wages than men.

Affirmative action programs are intended to promote women to overcome these disadvantages on labor markets. A very recent paper by Niederle, Segal and Vesterlund (2009) has investigated in a carefully designed laboratory experiment gender differences in competitiveness and how they are affected by an affirmative action program that guarantees a minimum percentage of women among the winners of a tournament with multiple winners. Niederle et al. (2009) have found that their affirmative action policy induces women to enter a competition (in a simple math task of adding two-digit numbers) much more often than in a control treatment without any policy to promote women. Interestingly, affirmative action had no negative effects on the efficiency of the tournament in selecting the best qualified candidates as winners since it encouraged in particular entry by high-performing women.²

In this paper we present an experimental study that adds to the literature on gender differences in competitiveness and the effects of policy interventions in two ways: First, we evaluate and compare three alternative types of policy interventions to promote women in

¹ Dreber, von Essen and Ranehill (2009), however, do not find men to increase their performance under competition more than women do in a study with Swedish school children. They argue that one reason for the lack of a gender difference might be the very egalitarian treatment of men and women in the Swedish society.

² Calsamiglia, Franke and Rey-Biel (2010) have reported that another form of affirmative action – extending preferential treatment to a group of disadvantaged subjects by giving them extra points in a sudoku-solving task – has had no negative effects either on the tournament's efficiency in selecting the best candidates as winners. Their study with 10-12 year old schoolchildren is not concerned with gender differences in competitiveness, though.

competitions: (A) *Quotas* that guarantee a certain minimum fraction of winners to be female, thus making the competition more gender-specific for women. For instance, many European parliaments have quotas on parliamentary seats that are reserved for women. (B) *Repetition of the Competition* if a critical number of female winners is not reached in the first attempt. For instance, in competitions for academic jobs (and more generally for jobs in the public sector) in Austria it is possible that the process of filling a vacant position is completely nullified and reset to the start if no woman is shortlisted for the position. This is essentially equivalent to repeating the competition then. (C) *Preferential treatment* of women by increasing their objective performance through adding a gender-specific bonus. Preferential treatment schemes are often encountered in practice, both in the public and the private sector, as a means to increase the participation of women in leading positions. A weak form of preferential treatment can be a tie-breaking rule that always favors women in case of equal performance or qualifications. In a stronger form, preferential treatment may imply discrimination against better-performing men. While policy intervention (A) has been studied by Niederle et al. (2009) and policy intervention (C) has been addressed by Calsamiglia et al. (2009) – however the latter not in the context of gender differences in competitiveness – we are able to compare in a unified framework the different types of interventions and their effects on behavior and on the efficiency in selecting the best candidates.

Second, we examine how the different policy interventions affect behavior *after* the competition. Competition within firms often means that one member of a workgroup receives a promotion, but that he or she still needs to work together with the other group members afterwards. Policy interventions in the spirit of affirmative action programs might backfire after a competition has been concluded by spoiling the willingness of losers to contribute to the group and coordinate efficiently in subsequent tasks. If losers perceive particular policy interventions as unfair they might be tempted to withhold effort in tasks where the whole group benefits from each single member's contribution or they might obstruct efficient coordination with others in coordination tasks. We study post-competition behavior in groups by letting our subjects perform two different tasks – a team task and a coordination game – after the competition has ended and its winners have been announced. The two tasks represent two different production functions in companies. In the team task each group member's output can be perfectly substituted by another member's output. This gives rise to problems of shirking in the spirit of Holmstrom (1982). Taking a minimum effort game as the coordination game implies a production function that determines a group's output by the minimum output in the group (for which reason the game is also known as weakest-link game). Coordination

problems abound in companies, such as showing up for meetings on time or not, sharing information with others or not, etc. (see Brandts and Cooper, 2006; Charness and Jackson, 2007). Our analysis of post-competition efficiency is, to our knowledge, the first attempt to investigate the possibility that policies aimed at promoting women in competition may impact on efficiency *after* the process of competition has been concluded.

Studying the effects of different policy interventions on competitive behavior of men and women – both in the course of a competition as well as after it – seems important for companies and politicians alike. Of course, companies and their human resource departments have a general interest in selecting the best candidates for a job, irrespective of gender. However, the gender composition of a company’s workforce can have implications for the company’s success (see, e.g., Weber and Zulehner, 2010). Therefore, companies may want to consider the gender of competitors in various ways, and also how different policies affect post-competition behavior. Likewise, politicians may want to provide an institutional and legal framework for a level-playing field of men and women on labor markets. This requires comparing different alternative measures and their effects on behavior and efficiency. In this paper we provide controlled laboratory evidence on the behavioral consequences and the efficiency of different intervention schemes to promote women in competitions.

In our experiment we had 360 participants. The experimental task was to add two-digit numbers without using calculators.³ The key parts of our design were a stage where participants could choose whether to compete against others with a tournament payment-scheme or opt for a piece-rate scheme, and two stages in which group members were exposed to incentives for shirking and problems of coordination in non-competitive tasks.

We find that all policy interventions to promote women in the competition (Quota, weak and strong Preferential Treatment, and Repetition of the Competition) have a positive impact on the proportion of women who choose to engage in competition, although the effect is not significant in case of Repetition of the Competition. Preferential treatment appears to generate the strongest incentives for women to compete. We estimate the impact of alternative policies on the actual and perceived probabilities of winning the tournament and – controlling for these probabilities – we find no significant residual gender effect on competition entry choices once any of the policy interventions applies. This evidence suggests that the policy interventions remove the gender differences in competition entry decisions (that are still

³ The reason for choosing this math task is not only that we follow the earlier literature (Niederle and Vesterlund, 2007; Niederle et al., 2009), but also because math test scores serve as a good predictor of future income (see, e.g., Murnane et al., 2000; Weinberger, 2001).

found in a control treatment) through affecting the (perceived and actual) probabilities of winning. In fact, we find that changes in the probabilities of winning account fully for the different entry choices across policies, so that there is no significant residual impact of policy on entry choices either. The only exception to the latter finding is the overreaction of women to the strong preferential treatment. The higher entry rates for women – combined with the fact that policy interventions have only a weak and insignificant negative effect on men’s entry choices – ensure that the different policy interventions do not have a negative effect on the competition’s efficiency in selecting the most qualified participants as winners. This result confirms under a unified framework the earlier findings by Niederle et al. (2009) and Calsamiglia et al. (2009) for particular policies. Studying different policies in the same experimental environment here reveals that all policy interventions lead to the same level of efficiency in selecting winners.

Our results with respect to post-competition behavior and efficiency show that implementing any of our four different policies does not have a negative effect on a group’s performance in a coordination game and a team task. In fact, in the treatments with a policy intervention we find a significantly *better* performance in the team task in comparison to a control treatment. This finding implies that policy interventions to promote women in competitions do not only increase women’s willingness to enter competition, but they may even be useful for a working group’s performance after the competition. We consider this a possible double dividend of affirmative action programs.

The rest of the paper is organized as follows: In Section 2 we introduce the experimental design. Section 3 presents the experimental results. Section 4 concludes the paper.

2. Experimental design

We used the experimental design of Niederle et al. (2009) as our point of departure and modified it along two dimensions, namely the number of policy interventions studied in different treatments and the introduction of post-competition stages. All experimental treatments described below share the following characteristics: At the beginning of the experiment subjects were randomly assigned into groups of six persons with three men and three women each. These groups stayed together for the whole duration of the experiment. All groups went through six different stages in the experiment. While subjects knew from the beginning that there would be six stages (see the experimental instructions in the appendix),

each stage was only introduced and explained after the previous one had been finished. The fact that there were fixed groups of six members was only announced after Stage 1. The experimental task in each of the stages 1 to 5 was adding five two-digit numbers in a limited time period of three minutes. Subjects were not allowed to use calculators, but could use scratch paper or do it off the top of their head. After each calculation subjects were informed whether their solution was correct or not, and then the next task was shown. The experimental task in stage 6 was a simple coordination game that is described more precisely as we explain the separate stages in what follows.

Stage 1 – Piece rate: Each subject received €0.50 for each calculation that was correctly solved within the limit of three minutes. This payment was independent of the other group members' performance.

Stage 2 – Tournament: In this stage group members had to compete against each other. The two members who had solved the most calculations correctly were paid €1.50 for each of them. The other four group members received nothing.

Stage 3 – Choice: Every group member could choose whether (s)he wanted to solve the calculations under a piece rate scheme (as in Stage 1) or a tournament scheme. If the tournament was chosen, then a subject's performance was compared to the other group members' performance in Stage 2⁴ and the rules for determining the winners differed across **treatments** as follows:

1. **Control treatment (CTR):** The winners were the two group members with the largest numbers of correct calculations, regardless of gender. Ties were broken randomly, as in all other treatments.
2. **Repetition of the competition (REP):** If there was not at least one woman among the two winners, the competition was repeated once. In this case, the rules of the repeated competition were the same as in the control treatment. Hence, gender played no longer any role if the competition had to be repeated.
3. **Minimum Quota (QUO):**⁵ Irrespective of the ordinal ranking of group members' performances, there had to be at least one woman among the two winners of the

⁴ Using the other group members' past performance has several advantages. First, tournament entry decisions do not depend on a subject's expectation about the other members' entry decisions, but only on the subject's beliefs about her/his own ability. Second, Stage 2 performances are competitive performances, and thus a subject competes against others when they were also exposed to a competitive payment scheme. Third, entering competition does not impose externalities on others. In principle, this means that Stage 3 is an individual decision making problem. Note that this scheme also implies that it is possible that all group members entering the competition in Stage 3 may win or all lose since they are competing against the others' performance in Stage 2.

⁵ This treatment is called the Affirmative Action (AA) treatment in Niederle et al. (2009).

tournament. This implied that the best performing woman was a winner for sure. The second winner could either be a man – if he performed better than all other men and better than the second-best woman – or a woman – if the second-best woman performed better than all men.

4. **Preferential Treatment 1 (PT1):** Each woman's performance was automatically increased by one unit (i.e., one correct calculation). Using the augmented scores for women, the rules of the control treatment applied, meaning that there were no further restrictions on the gender composition of the two winners.
5. **Preferential Treatment 2 (PT2):** Here each woman received two extra units as a head-start. Other than that the rules of **PT1** applied.

Note that all treatments were run in a between-subjects design. So each participant only experienced one particular payment scheme if the tournament was chosen in Stage 3. It is also important to stress that subjects did not receive any feedback on the outcome of the compulsory competition in Stage 2 or the optional competition in Stage 3 until the very end of the experiment. This was done in order to avoid that subjects condition their choices on previous outcomes of a competition. At the end of Stage 3 we elicited the beliefs of all subjects regarding their relative performance and their ranks in Stages 1 and 2. For each stage subjects had to indicate their expected rank within the whole group of six members, but also within their own gender only. Correct guesses were rewarded with €1 each, and the feedback was given also only at the end of the experiment.

Stage 4 – Compulsory tournament with policy intervention from Stage 3: In this stage the same treatments as in Stage 3 applied. However, subjects were forced to compete, and hence could not opt out as in Stage 3. As before, the two winners received €1.50 for each correct calculation. At the end of Stage 4 we informed subjects about the outcome of the competition in Stage 4 before moving on to the two non-competitive tasks in Stages 5 and 6. This means that each subject found out the gender of the two winners in the group (who were identified by means of an identification code that included an “F” for females and an “M” for males – see the instructions). In order to make winning and losing in Stage 4 more salient we gave each winner in Stage 4 an additional €5 as initial endowment in Stages 5 and 6, and each loser only €2. Our purpose was to introduce a clear distinction between winners and losers before starting with the post-competition stages.

Stage 5 – Team task: This task was identical in all treatments. Subjects had to add up two-digit numbers again. However, the payment scheme was such that each correct calculation was worth €0.50 for the group in total and then split equally among all group

members. Hence, the individual payoff from solving one calculation was only €0.083, rather than €0.50 as in the piece rate scheme, increasing the incentives to shirk and reduce effort in this task. We use the total group performance in Stage 5 in order to evaluate the impact of different policy interventions on a group's joint performance after a competition within the group has been concluded. While the monetary incentives are identical in Stage 5 across all treatments, the different experiences in Stage 4 might affect Stage 5-behavior in different ways.

Stage 6 – Coordination game: Each group member played the two-person coordination game illustrated in Table 1 with each of the other five group members. The so-called minimum effort game shown in Table 1 has 7 Nash-equilibria that are Pareto-ranked along the diagonal. Before picking a number from 1 to 7 a subject was informed about the gender of the other player (through the identification code) and whether this player had won or lost in the Stage 4-competition. With this information, each subject had to choose five times a number (which could be different, of course) for the interaction with each of the other group members. All decisions were made simultaneously and feedback was only given after Stage 6. Payments were determined by randomly forming three pairs in a group of six members and each subject was then paid according to the decisions made in his/her pair.

Table 1 about here

The experiment was run computerized with z-tree (Fischbacher, 2007) at the University of Innsbruck in October and November 2009. Using ORSEE (Greiner, 2004), we recruited 360 students from various academic backgrounds. We ran four sessions for each of our five treatments, with 18 subjects (i.e., three groups of six) in each session. This yields 12 groups with six members in each treatment. In half of the sessions in each treatment we let subjects play Stage 6 before Stage 5 in order to control for possible order effects of the team task and the coordination game. However, no order effects were found, allowing us to pool the data. In order to avoid wealth effects, one stage among Stages 1 to 4 and one stage among Stages 5 and 6 were randomly selected for payment at the end of the experiment. The flat payment of €5 (€2) for winners (losers) in Stage 4 was paid for sure. Each subject also received a show-up fee of €3. Our sessions lasted about 60 minutes and the average payoff per subject was €18.

3. Results

3.1. Performance in Stages 1 and 2

Figure 1 presents the average performance as number of correctly solved calculations across Stages 1 to 5. On average, men perform better than women. While the difference is not significant in the piece rate scheme of Stage 1 (6.44 vs. 6.02), men perform significantly better in the compulsory competition of Stage 2 (7.53 vs. 6.82; $p < 0.05$, two-sided Mann-Whitney U-test). The increase in performance from Stage 1 to Stage 2 is probably due to competition applying in Stage 2, although parts of this increase might also be driven by learning effects, since Figure 1 indicates an upward trend in performance. Note that the increase in performance is larger for men (+17%) than for women (+13%), but the difference is not significant. Table 2 presents the average performance (and its standard deviation) by gender for all five treatments. Controlling for gender, both in Stage 1 and Stage 2 there are no significant differences in performance between treatments (Kruskal-Wallis test). This was to be expected since Stages 1 and 2 are identical across treatments.

Figure 1 and Table 2 about here

3.2. Competition entry choices (Stage 3)

Figure 2 shows the relative frequency with which men and women choose to compete in Stage 3 (instead of choosing the piece rate scheme). Similar to the results reported in Niederle and Vesterlund (2007) and Niederle et al. (2009) we find that in the control treatment **CTR** men compete about twice as often as women. However, the four different policy interventions reduce, and even reverse, this gender gap. Overall, the relative frequency of women opting for competition is increasing from left to right in Figure 2, from 30.6% in **CTR** to 69.4% in **PT2**. Judging by a Chi²-test, the differences across treatments are highly significant ($p < 0.01$). Bilateral comparisons reveal that the frequency of competing women in **CTR** is significantly smaller than in **PT1** or **PT2** ($p < 0.05$), and weakly significantly smaller than in **QUO** ($p < 0.1$). Repetition of the competition (**REP**) is the only policy intervention that has no significant effect on women's entry choices.

Figure 2 about here

While the fraction of men choosing competition is generally decreasing from left to right in Figure 2, there is no significant difference across treatments (Chi²-test). Hence, the

main impact of the different policy interventions is on women's choices, not on the choices of men. In the following we analyze the reasons for the policies' consequences.

3.2.1. Probabilities of winning and actual versus optimal entry choices

First we calculate the probabilities of winning the tournament in Stage 3, conditional on gender and a subject's performance in Stage 2. For this purpose we draw 10,000 samples of groups of six members (3 men and 3 women; with replacement). The sampling is repeated 100 times and the average probability of winning is calculated for each member conditional on his/her performance in Stage 2, given the rule for determining the winners (**CTR**, **REP**, **QUO**, **PT1** or **PT2**). The resulting probabilities are shown in Table 3.

Table 3 and Table 4 about here

Given these probabilities, we can derive the payoff-maximizing entry choices and compare them with the actually observed entry choices. Assuming risk neutrality, entering the competition is payoff-maximizing if a subject's probability of winning the competition is larger than one third. This approach yields the number of subjects that should enter competition in Stage 3 as shown in Table 4 in rows "Payoff maximizing". Below these rows we report the actual number of men, respectively women, who entered the competition. We see from Table 4 that men always enter too often, and that the difference between the payoff-maximizing and the actual number of entrants is significant in three out of five treatments ($p < 0.05$, McNemar test). For women we do not find significant differences between the payoff-maximizing and the actual number of entrants, though. Rather, their choices are by and large consistent with conditional payoff maximization.

3.2.2. Econometric analysis of entry choices in the different treatments

In Table 5 we analyze the determinants of entry choices using a probit regression (marginal effects with robust standard errors; all findings are robust to individual random effects). The decision to enter the competition in Stage 3 (1 if a subject chooses competition, 0 otherwise) is regressed on a gender dummy, the four treatment dummies, the interactions between gender and treatments, and the following controls: A subject's competitive performance in Stage 2 is taken into account by variable *correct2*, measuring the number of correct answers in Stage 2. The variable *guesswin* is constructed from a subject's beliefs about his/her performance in Stage 2 and it takes on the value 1 if a subject has reported in his/her

beliefs either rank 1 or rank 2 for Stage 2, and 0 otherwise.⁶ The variable *probwin* represents a subject's expected probability of winning the tournament conditional on her Stage 2 performance, taken from Table 3. Hence, while *probwin* measures the ex ante probability of a subject winning the tournament in Stage 3, *guesswin* captures the perceived probability of winning (provided that subjects do not expect their relative performance to be different between Stages 2 and 3).

Table 5 about here

In column (1) of Table 5 we only control for the number of correct answers in Stage 2 (which has a positive effect on the likelihood to choose competition in Stage 3), but not for the probabilities of winning. We find a significant gender gap, indicated by the significantly negative female dummy. We also see that the interaction of *female* with treatment **PT2** is significantly positive, thus counterbalancing the negative gender gap in the control treatment **CTR**. This indicates an overshooting of women's entry decisions when the preferential treatment is strong. In column (2) we add *guesswin*, a subject's belief about winning the competition, and find that this renders the gender dummy *female* insignificant. The reason for that is the fact that women have significantly different beliefs about their rank than men. While 58.3% of men indicate in their beliefs a rank of 1 or 2 (meaning they expect to win), this proportion is only 38.9% for women ($p < 0.05$, Chi²-test). Hence, the negative gender gap in column (1) is mainly driven by different beliefs of men and women.⁷

In column (3) we add *probwin* as an explanatory variable. Since *probwin* and *correct2* are very highly correlated by construction of *probwin* (Spearman's rho = 0.92, $p < 0.01$), we exclude *correct2* from this specification, as we would run into problems of multicollinearity otherwise. Using *probwin* instead of *correct2* does not change any of the main results established so far. In columns (4) and (5) we break down the analysis of column (3) by gender, confirming that both male and female subjects respond to changes in the (actual and

⁶ If, instead of this binary measure of perceived beliefs, we use the ordinal rank in our regressions all our results stay qualitatively the same.

⁷ It may be interesting to note that gender differences in beliefs about winning the competition in Stage 2 are not driven by men being more overconfident. Recall that men have a significantly better performance in Stage 2 than women. Comparing their actual ranks with their expected ranks, we find an overconfidence of 0.53 ranks on average for men ($p < 0.01$, two-sided t-test), but we also find an overconfidence of 0.39 ranks on average for women ($p < 0.01$, two-sided t-test). The difference in overconfidence of men and women is not significant ($p > 0.2$).

perceived) probability of winning the competition.⁸ For both men and women there are no residual effects of any policy intervention, except for women entering significantly more often in **PT2** even after controlling for the probabilities of winning.

Overall, the results in columns (1) to (5) of Table 5 suggest that the (actual and perceived) probabilities of winning the competition can explain the differences in competition entry choices in Stage 3 so that – after controlling for these probabilities – there is no statistically significant residual gender gap in entry decisions (with an overreaction of female participants to the strong preferential treatment **PT2**, though). However, this overall result is driven by our treatments with a policy intervention and, therefore, masks the gender gap that still persists in the control treatment **CTR**. Using only data for **CTR**, we find that women choose to enter competition significantly less often than men (36.1% vs. 63.9%; $p < 0.01$, Chi²-test).⁹

3.3. Efficiency in selecting the best candidates as winners (Stages 3 and 4)

Policy interventions that promote the entry of women may have two opposing effects on the overall efficiency in selecting the best candidates as winners. On the one hand any policy intervention that gives an advantage to women (like in **QUO**, **PT1** and **PT2**) may yield efficiency losses by passing by better performing men for the sake of promoting women. We call this potential effect the selection effect of policy interventions. On the other hand policy interventions may induce more high-performing women to choose competition instead of going for the piece rate. What we call the entry effect of policy interventions may lead to efficiency gains. Whichever effect prevails is open to empirical examination.

Table 6 shows for each treatment the average Stage 1-performance of those subjects who have entered and won the competition in Stage 3. We use the winners' performances in Stage 1 as an appropriate measure of a subject's ability, because this performance is unaffected by competition and any policy intervention.¹⁰ The winners' average performance is better than in the control treatment **CTR** in three out of the four treatments with a policy intervention. A Kruskal-Wallis test shows that the differences across treatments are far from

⁸ Including the interactions of *probwin* and *guesswin* with the *female* dummy in column (3) would not lead to any notable changes in the results, as the interaction terms are insignificant and indicate that the effects of the probability of winning and of beliefs are not conditional on gender. This can also be seen from columns (4) and (5) by the small differences between the coefficients of *probwin* and *guesswin* for men and women.

⁹ We can replicate the significant gender gap in **CTR** even when we control for subjective beliefs and the actual probability of winning, by running a regression of the decision to enter competition in Stage 3 on *female*, *guesswin* and *correct2* and using only data for **CTR**: *female* is negative and significant in this regression ($p < 0.05$)

¹⁰ We obtain very similar results if we use performances in Stage 3 as our measure of ability (or performances in Stage 4; see Table 7 and the associated discussion).

being significant ($p > 0.7$), nor are any pairwise comparisons close to that ($p > 0.2$ for all pairwise comparisons, Mann-Whitney U-tests). These findings suggest that the entry effect and the selection effect balance out in the aggregate.

Table 6 and Figure 3 about here

Next, we have a closer look at the entry and the selection effect. Figure 3 plots in panel (a) the proportion of all subjects who choose to enter the competition in Stage 3, conditional on their performance in Stage 1. We classify Stage 1-performance as weak (less than four correct answers), intermediate (4 to 7 correct answers) or strong (more than 7 correct answers).¹¹ What we see from panel (a) is the fact that our four different policy interventions increase the likelihood of weak and strong performers entering the competition – always compared with **CTR** as benchmark – while they have no effect on the intermediate performers.

Panel (b) shows the same graph for women, indicating that the policy measures have a positive effect on competition entry for all types of female performers (with a negligible exception for intermediate performers in **REP**). The increase in competition entry by strong female performers shows the potential of policy interventions to increase efficiency (the entry effect).

Panel (c) presents the reaction of men to policy interventions. We see that in particular intermediate performers (with 4 to 7 correct answers) are discouraged from entering competition – especially in treatment **PT2** –, while strong male performers (with more than 7 correct answers) do not respond to policy interventions in a negative way, compared to **CTR** as benchmark.

Turning to the role of the selection effect for tournament efficiency, we present in Table 7 the average Stage 1-performance of the two winners in Stage 4 – where all participants had to compete. There is no significant difference across all treatments ($p > 0.30$, Kruskal-Wallis test), nor in pairwise comparisons to **CTR** ($p > 0.18$ in all cases, Mann-Whitney U-tests). This absence of an efficiency-decreasing selection effect is largely due to the fact that hardly any better-qualified men were passed by in the treatments with policy interventions. In treatment **REP** there was no single instance in which the competition had to be repeated in Stage 4, meaning that no better-qualified men could have been passed by. In

¹¹ This was done in order to have more observations in each category. The resulting patterns are qualitatively the same if we use a larger number of categories, or if we plot the proportions for each possible performance level.

treatment **QUO** there were three cases where a man performed better by one unit but lost to the best woman, and two cases where a man performed better by four units but lost to the best woman. Finally, in only one out of 12 groups in **PT1 (PT2)** a man performed better by one unit (two units) and lost.

Table 7 about here

3.4. Post-competition efficiency I: The team task (Stage 5)

Recall that we examine two different post-competition tasks: one where the group members' efforts are substitutes (in the team task in Stage 5) – which entails incentives for shirking – and one where group production has a weakest-link production function in the form of a minimum effort game (in Stage 6) – where the problem is how to coordinate efficiently. We start our analysis of post-competition efficiency with the team task.

Figure 4 about here

In Figure 4 we present the average total output as the sum of correct answers in a group in Stage 5 of the experiment. From an organizational point of view a higher output is more efficient. It is interesting to note that total output is higher in any of the treatments where group members had experienced a policy intervention in Stages 3 and 4 than in the control treatment **CTR**. Pooling all treatments with a policy intervention (i.e., **REP**, **QUO**, **PT1** and **PT2**) and testing against **CTR** yields a significantly higher output in the former set of treatments (48.4 vs. 44.5; $p < 0.05$, two-sided Mann-Whitney U-test).¹² More specifically, the highest output is achieved in treatment **PT2**, with a significant difference in a pairwise comparison to **CTR** ($p < 0.05$, Mann-Whitney test). We also find a weakly significant difference between **QUO** and **CTR** ($p < 0.1$).

We had conjectured that in particular men who had expected to win, but had actually lost a competition might withhold effort in Stage 5. In order to identify this set of men, we look at reported beliefs about Stage 2 rankings. Since the only difference between Stages 2 and 4 is policy (and given that there is no feedback on true relative performance before Stage 5), it is natural to assume that a man who thinks he has won in Stage 2 also thinks that he would win in Stage 4 unless he is disadvantaged by policy. Considering those men who

¹² Pooling of the treatments with a policy intervention is possible here since there are no significant Stage 5 differences among these treatments ($p > 0.4$, Kruskal-Wallis test).

expected to be first or second in their group in Stage 2, but who actually lost in Stage 4, leads to an interesting result: On average, male losers who expected to win have a *higher* Stage 5-performance in all policy intervention-treatments compared to **CTR** (7.83 vs. 6.64), and this difference is significant in the case of **PT2** ($p < 0.05$, two-sided Mann-Whitney U-test). The increase in male losers' performance could be part of an effort of those men who lost the tournament in Stage 4 to compensate for foregone payoffs (by working harder now) or demonstrate to their group that they would have deserved to win.

It is noteworthy that in the control treatment **CTR** men who thought they should have won, but actually lost, are slightly withholding effort, comparing performance in Stages 4 and 5 (7.0 vs. 6.64 correct answers). In contrast to this withholding of effort, the performance of men who expected to win, but lost, increases from Stage 4 and Stage 5 on average in all treatments where a policy intervention applied (7.35 vs. 7.83). This increase is, in fact, significant in **PT2** ($p < 0.05$, two-sided Wilcoxon signed ranks test).

For women we find no effects of policy interventions on their behavior in Stage 5. There is no significant difference across treatments in the Stage 5-performance of female winners, irrespective of whether they had expected to win or lose in Stage 2 (Kruskal-Wallis test). Similarly, we find no significant change in female losers' performance from Stage 4 to Stage 5, whether or not they had expected to win or lose. In sum, the evidence from the team task suggests that policy interventions may even be beneficial for team production, and mainly so because men increase their efforts.

3.5. Post-competition efficiency II: The minimum effort game (Stage 6)

Our second measure of post-tournament efficiency is a group's total payoff from the minimum effort game in Stage 6. We consider the sum of payoffs that all six group members would earn if every pairwise game were paid. The maximum total payoff would be €195, if all group members always chose an effort level of "7". The minimum total payoff would be €60. This would apply if in each pair there would be one player choosing "1", and the other "7", yielding an average payoff of €2 per subject and interaction. Hence, the total sum of payoffs in a group can be considered as an indicator of efficiency in a coordination task like the minimum effort game.

Figure 5 about here

Figure 5 shows the average total payoffs in a group in the different treatments. While they differ slightly – fluctuating around €150 – there is no significant difference across treatments ($p > 0.6$, Kruskal-Wallis test) and also no significant pairwise differences in comparison to the control treatment **CTR**.¹³ This leads us to conclude that, in the aggregate, introducing any of our policy measures does not entail efficiency losses in the minimum effort game.

Taken together, our results from Stages 4, 5 and 6 allow us to conclude that our policy interventions of minimum quotas, weak or strong preferential treatment, and repetition of the tournament do not entail any efficiency losses, neither in the selection of winners in the competition stage, nor in two different team tasks in post-competition stages. On the contrary, we have even found evidence that total team performance in the team task of Stage 5 is larger in groups that have experienced a policy intervention than in the control treatment.

4. Conclusion

Motivated by recent findings on the lower level of competitiveness of women (Gneezy et al., 2003, 2009; Niederle and Vesterlund, 2007; Niederle et al., 2009) we have examined in an experiment several types of policy interventions that are intended to promote women in competition by either giving them a head-start over men or by considering quotas when determining the winners of contests. The main targets of our study have been to examine the effects of the different intervention forms on (1) women's willingness to enter a competition and the efficiency in selecting the best candidates as winners *in* the competition, and (2) the behavior in two non-competitive group tasks *after* the competition.

In our experiment we have studied the effects of a weak and a strong form of preferential treatment (by giving women a smaller or larger head-start over men), the introduction of a minimum quota for female winners, and the repetition of the competition if a certain number of female winners is not reached. With respect to our first target we have found that both forms of preferential treatment (**PT1** and **PT2**) as well as the implementation of a minimum quota for female winners (**QUO**) have increased women's willingness to enter competition significantly. The strong form of preferential treatment has more than doubled the

¹³ There are no significant gender differences in effort choices across treatments either, and also no differences in effort choices of winners and losers of the competition in Stage 4. However, an interesting observation is that subjects are systematically more cooperative (i.e., choose higher numbers) when they are matched with someone of the opposite sex. This is true for both men and women, and in all five treatments ($p < 0.01$, Mann-Whitney tests).

entry rates of women. Hence, if an increase in the number of women entering competition is the overarching goal, then the strong form of preferential treatment is the preferred policy intervention. It seems that only the (in the German speaking area often applied) procedure of repeating a competition if women are not among the winners (for example by being shortlisted for positions in the public sector) does not have a significant impact on women's competition entry choices (**REP**). Overall, we can conclude, however, that women react systematically in their entry choices to policy interventions that are intended to promote them in tournaments. The main reasons for this reaction are the positive effects of these policy interventions on the actual and expected probabilities of winning the competition. Controlling for these probabilities we have seen that the policy interventions themselves have no significant effect on female entry choices, except for the strong preferential treatment **PT2** that leads to an excess entry of women. This is an indication that policy interventions might overshoot, thereby inducing inefficiently high entry rates of women into tournaments. In contrast to women, we have found that men react much less (negatively) – and insignificantly – to the different types of policy interventions. In a sense this finding is compatible with men enjoying competition – even if they are objectively disadvantaged.

As far as the efficiency in selecting the best qualified candidates as winners in the competition is concerned, it is comforting to note that none of the four policy interventions studied here had a negative effect in comparison to a control treatment. It seems that the higher entry rates of highly qualified women – combined with the fact that policy interventions have only a weak and insignificant negative effect on men's entry choices – is responsible for avoiding efficiency losses through measures that can potentially pass by better qualified men. None of our four different policy interventions has been found better or worse compared to each other, indicating that several different forms of affirmative action may yield roughly the same results with respect to efficiency in selecting winners. What we have found in a unified framework confirms the earlier findings of Niederle et al. (2009) and Calsamiglia et al. (2009) for particular forms of affirmative action.

Our second target has been to study how different forms of affirmative action influence post-competition behavior in groups when group members are expected to cooperate and coordinate efficiently. To the best of our knowledge, this issue has not been studied before under controlled conditions. We have chosen two different, non-competitive tasks to study post-competition behavior. Both tasks are representative for different types of production functions in teamwork. The team task treats each group member's effort as substitute for another member's effort, while the minimum effort game is a specific form of a

weakest-link production function. Studying both types of non-competitive tasks is motivated by the observation that while internal promotions in companies (through contests for better jobs) yield winners and losers, it is often necessary that winners and losers work together also after the contest. Different forms of policy interventions may have different effects on teamwork after the contest then. The fear that policy interventions might backfire on efficiency after the contest is unwarranted, though, as our findings show. In the coordination task we find no significant difference across all five treatments. However, in the team task we even find significantly higher output in the treatments with policy interventions than in the control treatment.

In sum, our findings seem to us like good news for companies and policy makers alike, provided that they have an interest in supporting women in competitive environments through various forms of affirmative action. While such interventions have – not unexpectedly – positive effects on the willingness of women to expose themselves to a competitive situation, they have – less expectedly – no negative effects on the efficiency of selecting the best candidates, they are not associated with efficiency losses from coordination failure, and they even have positive effects on productivity in a team task that had to be performed after the conclusion of the competition. Hence, affirmative action programs promoting women can have a double dividend, and this holds particularly true for the three forms of affirmative action (**QUO**, **PT1**, **PT2**) that have significantly positive effects on the decision of women to enter competition.

References

- Altonji Joseph, and Rebecca Blank.** 1999. "Race and Gender in the Labor Market." In: Ashenfelter, Orley, and David E. Card (Eds.), *Handbook of Labor Economics*, Vol. 3, Elsevier: 3143-3259.
- Bertrand, Marianne, and Kevin F. Hallock.** 2001. "The Gender Gap in Top Corporate Jobs." *Industrial and Labor Relations Review* 55(1): 3-21.
- Black, Sandra, and Philip E. Strahan.** 2001. "The Division of Spoils: Rent-sharing and Discrimination in a Regulated Industry." *American Economic Review* 91(4): 814-831.
- Blau, Francine, Marianne Ferber, and Anne Winkler.** 2010. "The Economics of Women, Men and Work." 6th edition. Englewood Cliffs, NJ: Prentice Hall.
- Blau, Francine, and Laurence Kahn.** 2000. "Gender Differences in Pay." *Journal of Economic Perspectives* 14(4): 75-99.
- Brandts, Jordi, and David J. Cooper.** 2006. "A Change Would Do You Good An Experimental Study on How to Overcome Coordination Failure in Organizations." *American Economic Review* 96(3): 669-693.
- Calsamiglia, Caterina, Jörg Franke, and Pedro Rey-Biel.** 2010. "The Incentive Effects of Affirmative Action in a Real-Effort Tournament." University Autònoma Barcelona, Working Paper.
- Charness, Gary, and Matthew O. Jackson.** 2007. "Group Play in Games and the Role of Consent in Network Formation." *Journal of Economic Theory* 136(1): 417-445.
- Croson, Rachel, and Uri Gneezy.** 2009. "Gender Differences in Preferences." *Journal of Economic Literature* 47(2): 448-474.
- Datta Gupta, Nabanita, Anders Poulsen, and Marie-Claire Villeval.** 2005. "Male and Female Competitive Behavior - Experimental evidence." IZA Bonn, Discussion Paper 1833.
- Dreber, Anna, Emma von Essen, and Eva Ranehill.** 2009. "Outrunning the Gender Gap – Boys and Girls Compete Equally." SSE/EFI Working Paper 709.
- Fischbacher, Urs.** 2007. "z-Tree: Zurich Toolbox for Readymade Economic Experiments". *Experimental Economics* 10(2): 171-178.
- Gneezy, Uri, Kenneth Leonard, and John List.** 2009. "Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society." *Econometrica* 77(5): 1637-1664.

- Gneezy, Uri, Muriel Niederle, and Aldo Rustichini.** 2003. "Performance in Competitive Environments: Gender Differences." *Quarterly Journal of Economics* 118(3): 1049-1074.
- Gneezy, Uri, and Aldo Rustichini.** 2004. "Gender and Competition at a Young Age." *American Economic Review, Papers and Proceedings* 94(2): 377-381.
- Goldin, Claudia, and Cecilia Rouse.** 2000. "Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians." *American Economic Review* 90(4): 715-741.
- Greiner, Ben.** 2004. "An Online Recruitment System for Economic Experiments." In: Kremer, Kurt and Volker Macho (Eds.): *Forschung und wissenschaftliches Rechnen 2003*. GWDG Bericht 63. Gesellschaft für Wissenschaftliche Datenverarbeitung, Göttingen: 79-93.
- Holmstrom, Bengt.** 1982. "Moral Hazard in Teams." *Bell Journal of Economics* 13(2), 324-340.
- Murnane, Richard J, John B. Willett, Yves Buhaldeborde, and John H. Tyler.** 2000. "How Important Are the Cognitive Skills of Teenagers in Predicting Subsequent Earnings?" *Journal of Policy Analysis and Management* 19(4): 547-568.
- Niederle, Muriel, Carmit Segal, and Lise Vesterlund.** 2009. "How Costly Is Diversity? Affirmative Action in Light of Gender Differences in Competitiveness." National Bureau of Economic Research Working Paper 13923.
- Niederle, Muriel, and Lise Vesterlund.** 2007. "Do Women Shy Away from Competition? Do Men Compete too Much?" *Quarterly Journal of Economics* 122(3): 1067-1101.
- Niederle, Muriel, and Lise Vesterlund.** 2010. "Explaining the Gender Gap in Math Test Scores: The Role of Competition." *Journal of Economic Perspectives*, forthcoming.
- Weber, Andrea, and Christine Zulehner, C.** 2010. "Female Hires and the Success of Start-up Firms." *American Economic Review, Papers and Proceedings*, forthcoming.
- Weichselbaumer, Doris, and Rudolf Winter-Ebmer.** 2007. "The Effects of Competition and Equal Treatment Laws on Gender Wage Differentials ." *Economic Policy* 22(50): 235-287.
- Weinberger, Catherine J.** 2001. "Is Teaching More Girls More Math the Key to Higher Wages?" In: *Squaring Up: Policy Strategies to Raise Women's Income in the U.S.*, edited by King, M. C. University of Michigan Press.
- Wozniak, David, William T. Harbaugh, and Ulrich Mayr.** 2010. "Choices about Competition: Differences by Gender and Hormonal Fluctuations, and the Role of Relative Performance Feedback." University of Oregon, Working Paper.

Tables and Figures

Table 1. Payoff matrix in the minimum effort game (Stage 6).

		Minimum of the two numbers in a pair (including your number)						
		7	6	5	4	3	2	1
Your number	7	€6.50	€5.50	€4.50	€3.50	€2.50	€1.50	€0.50
	6		€6.00	€5.00	€4.00	€3.00	€2.00	€1.00
	5			€5.50	€4.50	€3.50	€2.50	€1.50
	4				€5.00	€4.00	€3.00	€2.00
	3					€4.50	€3.50	€2.50
	2						€4.00	€3.00
	1							€3.50

Table 2. Average performance (# correct answers) and standard deviation by stage, treatment and gender.

	CTR		REP		QUO		PT1		PT2	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
<i>Stage 1</i>	6.39 (2.84)	5.81 (2.42)	6.31 (2.48)	5.67 (2.35)	6.89 (3.22)	6.08 (2.39)	6.31 (3.00)	6.64 (2.87)	6.28 (2.63)	5.89 (1.98)
<i>Stage 2</i>	7.33 (2.99)	6.56 (2.48)	7.41 (2.75)	6.81 (2.49)	8.11 (3.09)	6.44 (2.80)	6.86 (3.39)	7.61 (3.08)	7.92 (2.47)	6.67 (2.32)
<i>Stage 3</i>	7.50 (3.42)	7.41 (2.38)	7.69 (2.86)	7.53 (2.41)	8.03 (3.51)	7.39 (2.64)	7.58 (3.28)	7.81 (3.54)	8.64 (3.08)	6.67 (2.81)
<i>Stage 4</i>	7.86 (2.94)	7.31 (2.69)	8.44 (2.78)	7.58 (2.05)	8.06 (3.35)	7.42 (2.95)	7.69 (3.79)	8.00 (3.32)	8.81 (2.66)	7.47 (2.34)
<i>Stage 5</i>	7.61 (3.13)	7.22 (2.87)	7.94 (2.94)	7.94 (2.28)	8.50 (3.61)	7.67 (2.69)	7.86 (3.38)	7.83 (3.02)	9.19 (3.15)	7.56 (2.62)

Table 3. Probability of winning the tournament in Stage 3, conditional on Stage 2 performance.

Stage 2 performance		5	6	7	8	9	10	11	12	13	14
(# correct answers)											
CTR	Men	0.01	0.07	0.24	0.48	0.68	0.82	0.91	0.97	0.99	1.00
	Women	0.01	0.06	0.22	0.45	0.66	0.79	0.89	0.96	0.99	1.00
REP	Men	0.01	0.07	0.21	0.47	0.66	0.81	0.90	0.97	0.99	1.00
	Women	0.01	0.07	0.28	0.53	0.74	0.86	0.94	0.98	1.00	1.00
QUO	Men	0.01	0.04	0.14	0.31	0.48	0.61	0.72	0.82	0.91	0.96
	Women	0.07	0.19	0.37	0.58	0.75	0.85	0.91	0.97	0.99	1.00
PT1	Men	0.00	0.03	0.13	0.33	0.56	0.74	0.85	0.94	0.96	1.00
	Women	0.03	0.14	0.35	0.57	0.74	0.85	0.94	0.98	1.00	1.00
PT2	Men	0.00	0.01	0.05	0.19	0.40	0.62	0.78	0.89	0.96	0.99
	Women	0.08	0.25	0.46	0.66	0.80	0.90	0.97	0.99	1.00	1.00

Table 4. Number of subjects (out of 36) entering competition in Stage 3.

	CTR	REP	QUO	PT1	PT2
Payoff maximizing (men)	13	15	17	10	13
Actual (men)	23	24	22	20	18
<i>p-value (McNemar test)</i>	0.01	0.03	0.20	0.00	0.13
Payoff maximizing (women)	12	15	13	23	19
Actual (women)	11	14	19	21	25
<i>p-value (McNemar test)</i>	0.80	0.76	0.08	0.59	0.11

Table 5. Probit regressions with *entry choice* into competition as dependent variable (Stage 3).

	(1) All	(2) All	(3) All	(4) Men	(5) Women
<i>female</i>	-0.308*	-0.185	-0.186		
	(0,115)	(0,133)	(0,133)		
<i>female*REP</i>	0.048	-0.116	-0.132		
	(0,171)	(0,184)	(0,183)		
<i>female*QUO</i>	0.278	0.238	0.211		
	(0,130)	(0,148)	(0,154)		
<i>female*PT1</i>	0.275	0.188	0.159		
	(0,131)	(0,154)	(0,160)		
<i>female*PT2</i>	0.435 **	0.411 **	0.376 **		
	(0,073)	(0,085)	(0,098)		
<i>REP</i>	0.031	0.098	0.101	0.100	-0.025
	(0,122)	(0,125)	(0,125)	(0,120)	(0,130)
<i>QUO</i>	-0.065	-0.067	-0.051	-0.051	0.175
	(0,126)	(0,134)	(0,134)	(0,132)	(0,124)
<i>PT1</i>	-0.062	-0.005	0.012	0.012	0.177
	(0,124)	(0,124)	(0,123)	(0,120)	(0,124)
<i>PT2</i>	-0.171	-0.157	-0.123	-0.122	0.326 **
	(0,118)	(0,130)	(0,130)	(0,131)	(0,106)
<i>correct2</i>	0.047 **	0.028 *			
	(0,011)	(0,011)			
<i>guesswin</i>		0.348 **	0.346 **	0.358 **	0.332 **
		(0,055)	(0,055)	(0,077)	(0,078)
<i>probwin</i>			0.266 **	0.243 *	0.286 *
			(0,089)	(0,125)	(0,126)
Pseudo R ²	0.086	0.158	0.162	0.142	0.172
Prob>chi2	0.000	0.000	0.000	0.000	0.000
N	360	360	360	180	180

* (**) ... significant at the 5% (1%) level. The table reports marginal effects.

Table 6. Tournament efficiency – Average Stage 1-performance of the winners in Stage 3, by treatment.

CTR	7.79
REP	8.17
QUO	8.56
PT1	8.38
PT2	7.52

Table 7. Selection effect – Average Stage 1-performance of the winners in Stage 4, by treatment.

CTR	8.13
REP	7.42
QUO	8.33
PT1	9.13
PT2	7.50

Figure 1. Average performance by gender across all treatments (# correct answers).

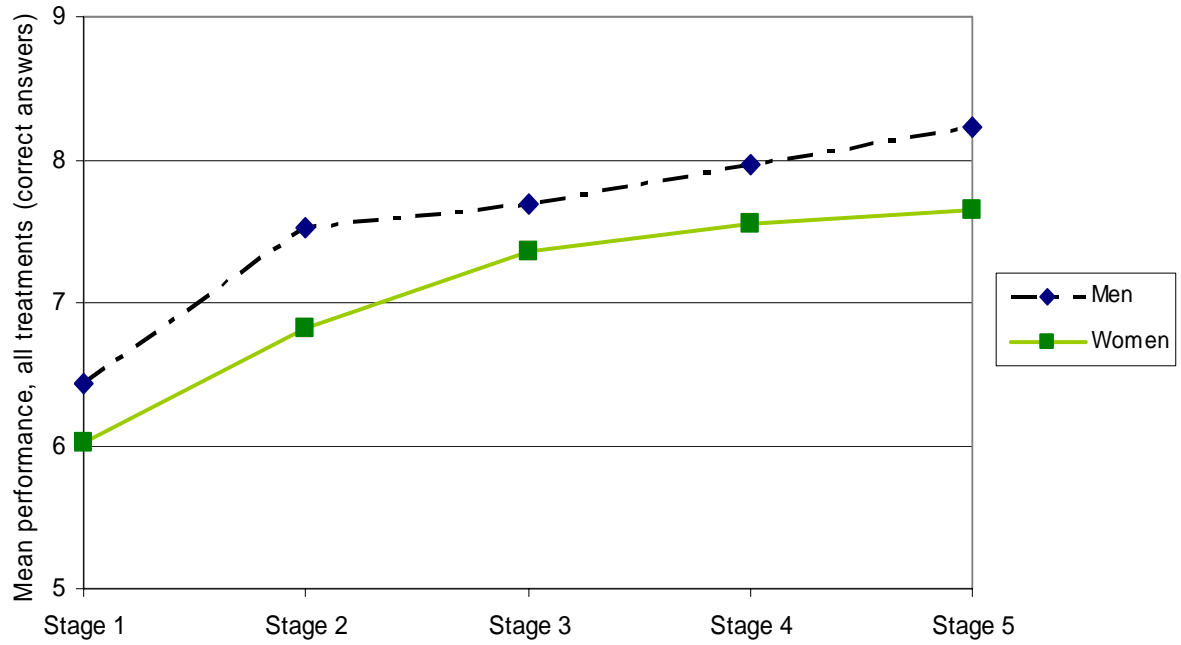


Figure 2. Relative fraction of subjects choosing competition in Stage 3.

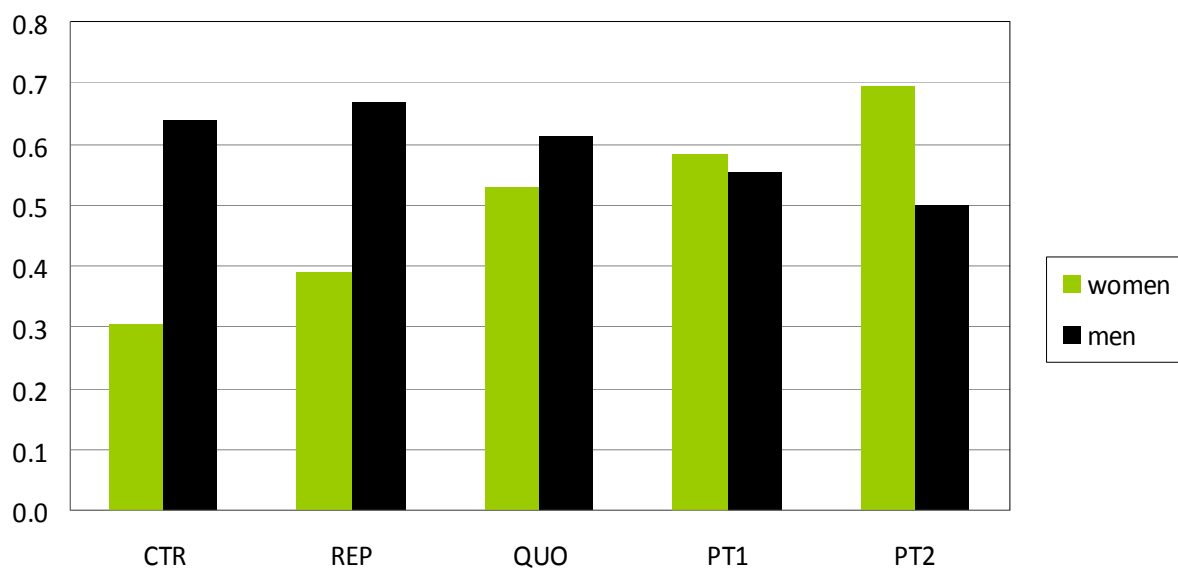


Figure 3. Relative frequency of entering competition in Stage 3, conditional on performance in Stage 1.

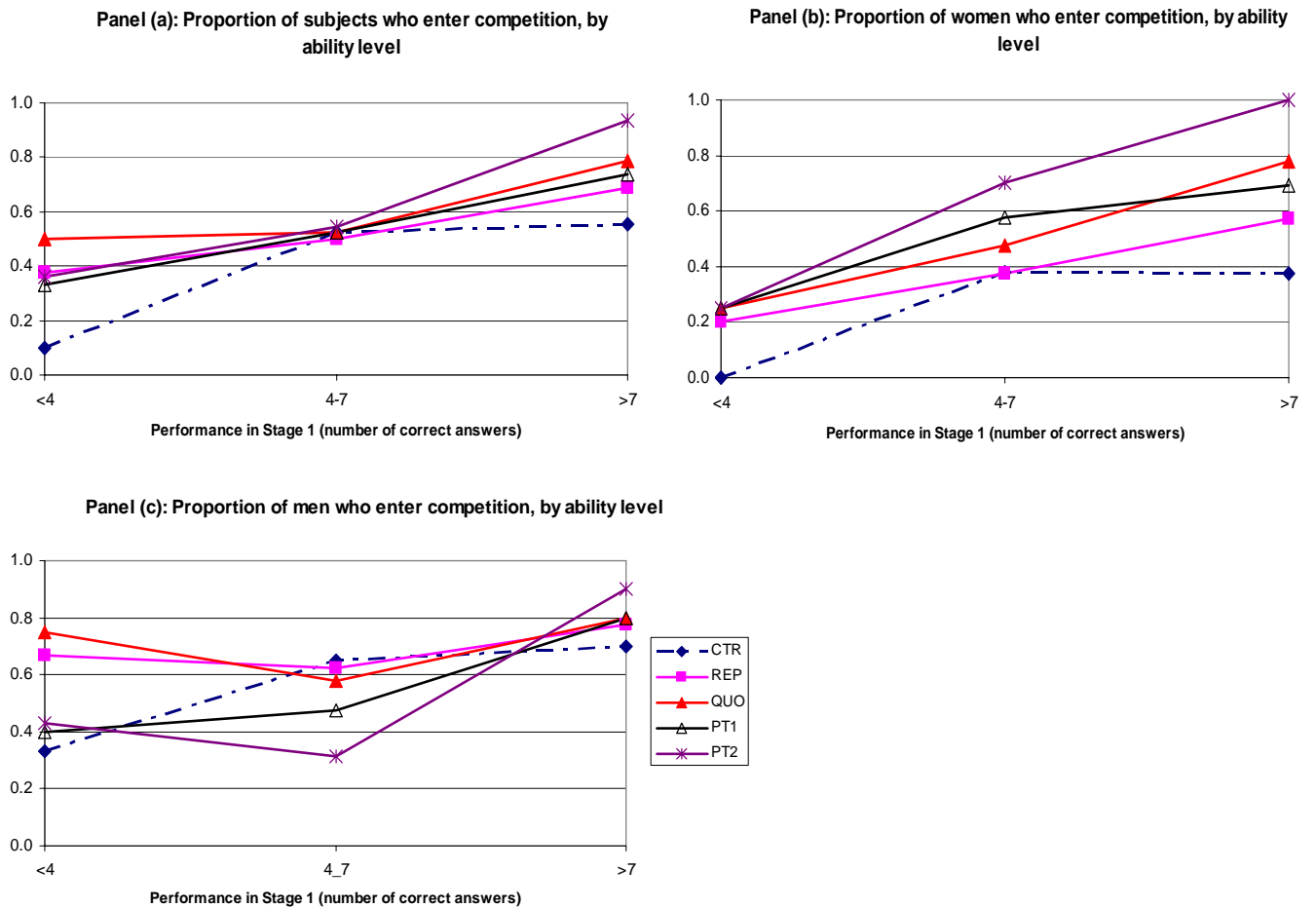


Figure 4. Average group performance in Stage 5.

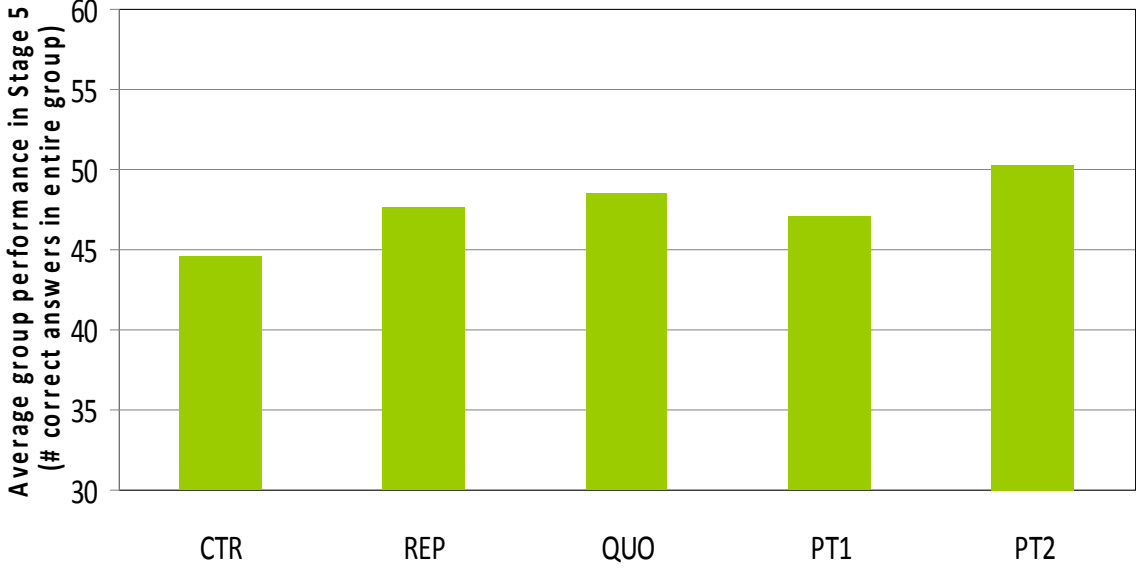
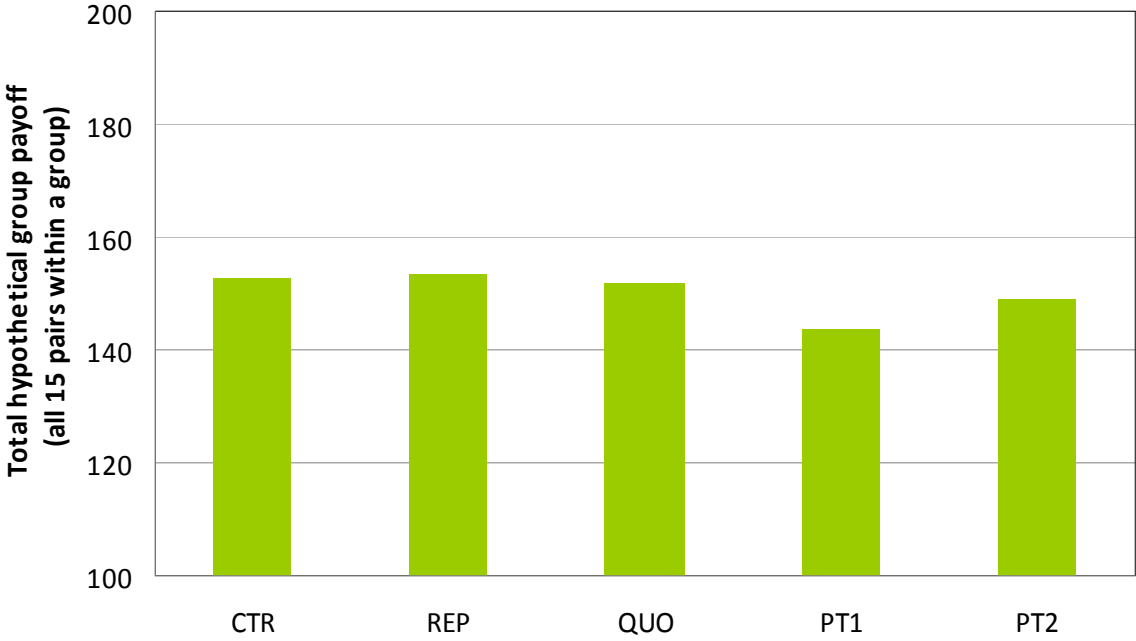


Figure 5. Total hypothetical group payoff in the minimum effort game of Stage 6.



Appendix: Experimental Instructions (*not intended for publication*)

[General instructions, at start of session]

Welcome to an experiment on decision making. We thank you for your participation!

During the experiment, you and the other participants will be asked to make certain decisions. Your own decisions as well as the decisions of the other participants will determine your payment from the experiment, according to the rules that will be described in what follows.

The experiment will be conducted on the computer. You make your decisions on the screen. All decisions and answers will remain confidential and anonymous.

The experiment consists of 6 stages. One of the first four stages (1-4) and one of the last two stages (5-6) will be randomly selected for your payment: At the end of the experiment we will use a lottery wheel to determine which stages will be relevant for your payment. Your total earnings from the experiment will be the sum of your payments for the randomly selected stages, plus a show up fee of €3.

You will receive instructions for each of the six stages, one after the other. We will read the instructions aloud and then give you time for questions. Please do not hesitate to ask questions if anything is not clear.

Please do not talk to each other during the experiment. If you have any questions, please raise your hand.

Stage 1: Piece rate

Your task in stage 1 is to solve correctly as many addition exercises as possible. To be more precise, you will have 3 minutes's time in order to solve as many additions of five randomly selected two-digit numbers as possible, by entering the sum of the five numbers. You are not allowed to use calculators but you can write down the numbers and use the provided scribbling paper for your calculations. You enter an answer by clicking with the mouse on the "Confirm" button. When you enter an answer, you immediately find out on the screen whether it was correct or not.

If stage 1 is the stage selected for payment (among stages 1-4), then you will receive **€0.50 (i.e., 50 cent) for each correct answer** that you entered within the 3 minutes. Your payment is not reduced when you enter a wrong answer. From now on, we call this method of payment the Piece-rate payment.

Directly before the start of this stage you will be given one minute in order to familiarize yourselves with the screen: During this time you can solve addition exercises, which do not count for the experiment. Afterwards, stage 1 will begin.

Stage 2: Tournament

As in stage 1, you will have 3 minutes' time in order to solve correctly as many addition exercises as possible. However, your payment in this stage depends on your performance relative to the performance of a group of participants.

Allocation in groups: Each group consists of 6 participants, 3 of whom are men and 3 are women. Groups are randomly formed at the beginning of this stage and **each participant stays in the same group until the end of the experiment.**

Each group member receives an identification code. All members keep their identification code for the entire experiment. **The 3 women receive randomly the identification code F1, F2 or F3. The 3 men receive randomly the identification code M1, M2 or M3.** You will not find out the identity of the other participants in your group during or after the experiment, so that all decisions remain anonymous.

If stage 2 is the stage selected for payment (among stages 1-4), then your payment depends on how many additions you have solved correctly in comparison with the other five participants in your group. The two group members who have entered the most correct answers are the two winners of the tournament. The two winners receive **€1.50 per correct answer** each, while the other four members **do not receive any payment**. In case of a tie, the ranking among the members with equal performances is determined randomly. From now on, we call this method of payment the Tournament payment.

You will not be informed about the outcome of the tournament until the end of the experiment.

Tournament, format B [*Preferential treatment only – PT2 {PT1}*]

Before the start of the next stage we explain the rules of the tournament, format B, from now on called "Tournament-B".

The only difference between the Tournament-B, and the tournament in stage 2 is the following: In the Tournament-B, the number of every woman's correctly solved exercises is automatically increased by 2. This means for example that, if a woman has entered 8 correct answers, her performance in the Tournament-B counts as 10 correct answers. In each group, all 3 women receive two *{one}* additional points each, while the 3 men receive no additional points. As in stage 2, the two winners of the Tournament-B are then the two group members with the best performances (taking the additional points for women into account). In case of a tie, the ranking among the members with equal performances is again determined randomly.

Tournament, format B [*Quotas only - QUO*]

Before the start of the next stage we explain the rules of the tournament, format B, from now on called "Tournament-B".

In the Tournament-B, the two winners are determined as follows. In each group, one of the two winners is in any case the woman with the best performance (of all three women). The other winner is the group member with the best performance among the remaining members (i.e., excluding the best-performing woman).

We now give a concrete example, in order to illustrate the way that the winners are determined in the Tournament-B. We order the six group members according to their performance within each gender, so that fA is the woman with the best performance, fB is the woman with the second-best performance and fC the woman with the third-best performance. In the same way, mA is the man with the best performance, mB is the man with the second-best performance, and mC the man with the third-best performance. The woman with the best performance, fA , is definitely one of two winners in a Tournament-B. In order to determine the second winner, we must find out who is the person with the best performance among the remaining five group members (besides fA). Since there is only one more winner, this can either be fB or mA , depending on their performance.

To sum up: A woman wins a Tournament-B if she has the best performance among all women or if she is one of the two persons with the highest performance within her group. A man wins a Tournament-B if he is the man with the best performance and at the same time one of the two persons with the highest performance within his group. So there is at least one woman and at most one man as winners in a Tournament-B.

Tournament, format B [*Repetition only - REP*]

Before the start of the next stage we explain the rules of the tournament, format B, from now on called “Tournament-B”.

The only difference between the Tournament-B, and the tournament in stage 2 is the following: If both winners in the Tournament-B are male, then the tournament is repeated **once**. The outcome of the second tournament only determines then, who the two winners are.

Thus, if there is at least one woman among the two winners of a Tournament-B, then the outcome is final. If both winners are men, then the tournament is repeated once. The repeated tournament is like the first one, i.e., the winners are the two group members who have entered the most correct answers. The outcome of the repeated tournament is final, even if for example both winners are men.

Stage 3: Choice between Piece-rate payment and Tournament-B payment

[In control treatment: *Replace Tournament-B with Tournament*]

As in stages 1 and 2, you will have 3 minutes’ time in order to solve correctly as many addition exercises as possible. However, you must now choose yourself your preferred payment method for your performance in stage 3. You can either choose the Piece-rate payment (as in stage 1) or the Tournament-B payment.

If stage 2 is the stage selected for payment (among stages 1-4), then your payment is determined as follows:

- If you choose the **Piece-rate payment**, then you will receive **€0.50 per correct answer**
- If you choose the **Tournament-B payment**, then your performance in stage 3 will be evaluated in comparison to the performance of the other five group members **in stage 2 (Tournament)**. As a reminder: That is the stage that you have just completed. If you enter more correct answers than four of your group members did in stage 2, then you will receive **€1.50 per correct answer** (i.e., 3 times the Piece-rate payment). In other words, only one member of your group can have a stage 2- performance which is higher than your stage 3- performance, otherwise you receive **no payment** for this stage. [**Preferential treatment only:** *As we have already explained, the performances of all women in the Tournament-B are automatically increased by two correctly solved exercises. The additional points are taken into account for the determination of the two winners, but not for the payment. This is true for the performances of women in stage 3, but also for the performances from stage 2 which are used for comparison.*] [**Repetition only:** *As we have already explained, the Tournament-B is repeated once if both winners are men. This means that those persons who choose the Tournament-B payment may have to participate in a tournament twice.*] [**Quotas only:** *The instructions for women are: You receive € 1.50 per correct answer if you have a better stage 3- performance than (i) the other two women in your group in stage 2, or (ii) four members of your group in stage 2. The instructions for men are: You receive € 1.50 per correct answer if you have a better performance than (i) the other two men in your group in stage 2, and (ii) four members of your group in stage 2.*]

In case of a tie, the ranking among the members with equal performances is again determined randomly.

The group composition (with 3 men and 3 women) is as in stage 2. If you choose the Tournament-B payment, you will not be informed about the outcome of the tournament until the end of the experiment.

On the next screen you will be asked whether you want to choose the Piece-rate payment or the Tournament-B payment for your performance in stage 3. Afterwards you will have 3 minutes in order to calculate the sums of the two-digit numbers.

Stage 4: Tournament-B

As before, you will have 3 minutes' time in order to solve correctly as many addition exercises as possible. The group composition (with 3 men and 3 women) is the same as before.

[In treatment PT2 {PT1}: The rules of the Tournament-B in this stage are as we have already described them in stage 3: The number of every woman's correctly solved exercises is automatically increased by 2 {1}. Hence, the only difference compared with stage 3 is that you now do not have a choice between the Piece-rate and the Tournament-B payment: All group members compete simultaneously in the Tournament-B.]

[In treatment QUO: The rules of the Tournament-B in this stage are as we have already described them in stage 3: In each group, one of the two winners is in any case the woman with the best performance (among all three women). Hence, the only difference compared with stage 3 is that you

now do not have a choice between the Piece-rate and the Tournament-B payment: All group members compete simultaneously in the Tournament-B.]

[In treatment REP: *The rules of the Tournament-B in this stage are as we have already described them in stage 3: If both winners in the Tournament-B are male, then the tournament is repeated once. Hence, the only difference compared with stage 3 is that you now do not have a choice between the Piece-rate and the Tournament-B payment: All group members compete simultaneously in the Tournament-B.]*

[In control treatment: *The tournament in this stage is exactly as in stage 2: The winners are the two group members who enter the most correct answers.]*

If stage 4 is the stage selected for payment (among stages 1-4), then the two winners receive **€1.50 per correct answer** each, while the other four members of the group receive **no payment**. In case of a tie, the ranking among the members with equal performances is again determined randomly.

Furthermore, the tournament in this stage also serves the purpose of determining the initial endowment of every group member in the next two stages (5 and 6). This is done as follows: The two winners of stage 4 receive then in stages 5 and 6 an initial endowment of **€5.00** each; the other four group members (i.e., the non-winners) receive an initial endowment of **€2.00** each in each of the next two stages. This endowment is only paid out for that stage (between stages 5 and 6) which is randomly selected for payment at the end of the experiment.

At the end of this stage you will be informed about the outcome of the tournament and thereby about your initial endowment in the next two stages. This means that all members in each group will find out who the two winners are (identified by means of their identification codes).

Stage 5

In stage 5, the two winners of the tournament in stage 4 have an endowment of €5; the non-winners have an endowment of €2.

As in stages 1-4, you will have 3 minutes' time in order to solve correctly as many addition exercises as possible. The group composition (with 3 men and 3 women) is the same as before. However, your payment for this stage depends on your performance as well as on the total performance of all other members in your group.

To be more precise, your payment is determined as follows: You receive **8.33 euro cent** for each correct answer that a member of your group has entered in the 3 minutes. This means that each correct answer is worth **€0.50 for the entire group** (i.e., all 6 members). It also means that all members of a group receive the same payment in this stage, and **this depends only on the total performance of the group, i.e., on the sum of all correct answers of the 6 group members.**

If stage 5 is the stage selected for payment (between stages 5-6), then your payment is the sum of your endowment and your payoff from the performance of your group.

At the end of the experiment you will be informed about the total performance of your group in this stage.

Stage 6

In stage 6, the two winners of the tournament in stage 4 have an endowment of €5; the non-winners have an endowment of €2.

In stage 6, pairs are formed within each group. Each pair consists of two persons. In this way, every participant belongs to five pairs, i.e., each time with one of the other five members of his (her) group.

In this stage you must make a decision by selecting one of the numbers 1, 2, 3, 4, 5, 6, or 7. Your payment depends on your own number as well as on the number selected by the other person that you are paired with. This is done as follows: In the table below you see in rows (left) the number that you have selected. In the columns (up) you see the minimum between your number and the number selected by the other person in your pair. The amounts in the cells are your profit, dependent on your number and on the minimum of the two numbers in your pair.

		Minimum of the two numbers in a pair						
		7	6	5	4	3	2	1
Your number	7	€6.50	€5.50	€4.50	€3.50	€2.50	€1.50	€0.50
	6		€6.00	€5.00	€4.00	€3.00	€2.00	€1.00
	5			€5.50	€4.50	€3.50	€2.50	€1.50
	4				€5.00	€4.00	€3.00	€2.00
	3					€4.50	€3.50	€2.50
	2						€4.00	€3.00
	1							€3.50

Example 1: Suppose you select the number 6 and the other person in your pair selects the number 4. Then 4 is the minimum between these two numbers. From the table it follows that your profit is €4, while the profit of the other person is €5.

Example 2: Suppose you select the number 2 and the other person in your pair selects the number 5. Then 2 is the minimum between these two numbers. From the table it follows that your profit is €4, while the profit of the other person is €2.50.

Example 3: Suppose you select the number 3 and the other person in your pair also selects the number 3. Then 3 also is the minimum between these two numbers. From the table it follows that your profit is €4.50, and the profit of the other person is also €4.50.

The concrete procedure in this stage is as follows: You will be asked on the screen to make five decisions, one for your interaction with each of the other members of your group. For each decision you will be informed on-screen with which group member you are playing (identified by means of the identification code). You will, however, not find out which number (s)he has chosen, before you make your own choice. In other words, the two persons in each pair make their decisions simultaneously.

For your payment in this stage one of the five pairs, to which you belong, will be randomly selected. If stage 6 is the stage selected for payment (between stages 5-6), then your payment is the sum of your endowment and your payoff based on the above table.

At the end of the experiment you will be informed about the decision of the other person in the pair that is relevant for your payment, and thereby about your payoff from this stage. However, you will not find out anything about the decisions in the other pairs (to which you do not belong).