CENTRE FOR DECISION RESEARCH & EXPERIMENTAL ECONOMICS

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Discussion Paper No. 2016-20

Felix Kölle

December 2016

Affirmative Action and Team Performance

CeDEx Discussion Paper Series
ISSN 1749 - 3293

CENTRE FOR DECISION RESEARCH & EXPERIMENTAL ECONOMICS

The Centre for Decision Research and Experimental Economics was founded in 2000, and is based in the School of Economics at the University of Nottingham.

The focus for the Centre is research into individual and strategic decision-making using a combination of theoretical and experimental methods. On the theory side, members of the Centre investigate individual choice under uncertainty, cooperative and non-cooperative game theory, as well as theories of psychology, bounded rationality and evolutionary game theory. Members of the Centre have applied experimental methods in the fields of public economics, individual choice under risk and uncertainty, strategic interaction, and the performance of auctions, markets and other economic institutions. Much of the Centre's research involves collaborative projects with researchers from other departments in the UK and overseas.

Please visit http://www.nottingham.ac.uk/cedex for more information about the Centre or contact

Suzanne Robey
Centre for Decision Research and Experimental Economics
School of Economics
University of Nottingham
University Park
Nottingham
NG7 2RD
Tel: +44 (0)115 95 14763
Fax: +44 (0) 115 95 14159
suzanne.robey@nottingham.ac.uk

The full list of CeDEx Discussion Papers is available at

http://www.nottingham.ac.uk/cedex/publications/discussion-papers/index.aspx

Affirmative Action and Team Performance

Felix Kölle[1]

University of Cologne

December 16, 2016

Abstract: We experimentally investigate spillover effects of affirmative action policies in tournaments on subsequent team performance and the willingness to work in teams. In three different team environments, we find that such policies in form of gender quotas do not harm performance and cooperation within teams, and do not weaken people's willingness to work in teams. Our results, thus, provide further evidence that gender quotas can have the desired effect of promoting women without harming efficiency.

Keywords: Affirmative action; teams; gender; performance; cooperation; selection; experiment

JEL: C91, C92, D03

[1] Faculty of Management, Economics and Social Sciences, University of Cologne, Germany. Tel.: +49 221 4705317. E-mail address: felix.koelle@uni-koeln.de. Financial support from the Deutsche Forschungsgemeinschaft (DFG) and the Austrian Science Fund (Project P22722-G11) is gratefully acknowledged.

1. Introduction

Gender differences in labor market outcomes are a prevalent and hotly debated topic in society and public policy. Despite several improvements over the past decades, there are still substantial gender differences with respect to, e.g., wages, career advancement, and the representation in executive positions (e.g., Blau and Khan, 2000; Weichselbaumer and Winter-Ebmer, 2007; Blau et al., 2013). One finding that has been put forward as a potential explanation for these differences is that women tend to shy away from competition both in the lab and in the field (e.g., Gneezy et al., 2003; Niederle and Vesterlund, 2007; Croson and Gneezy, 2009; Booth and Nolen, 2012; Datta Gupta et al., 2013; Brandts et al., 2014; Buser et al. 2014; Sutter and Rützler, 2015).[2] Affirmative action policies intended to promote women by encouraging them to compete have become a popular tool aimed at diminishing this imbalance. Recent experimental studies examined the efficacy of such programs, showing that they indeed have the desired effect of reducing the gender gap in the willingness to compete without harming efficiency in tournaments (e.g., Balafoutas and Sutter, 2012; Villeval, 2012; Niederle et al., 2013; Calsamiglia et al., 2013).

Much less is known, however, about potential detrimental effects of affirmative action policies in different environments. The implementation of quotas in tournaments, for example, might create negative spillover effects on subsequent interactions if perceived as unfair. Such situations are relevant, for instance, in firms when members of the same team or department compete for a promotion, and afterwards have to collaborate with the very same people again. Discontent about (potentially unfair) promotional decisions that are not solely based on performance might then harm work morale. This, in turn, might negatively affect team performance either in form of reduced effort, or in form of increased adverse behavior. Furthermore, this might also reduce the willingness to work in teams, which is important because firms increasingly rely on teams and the use of team incentives when organizing their workforce (see e.g., Hamilton et al., 2003; Lazear and Shaw, 2007).

In this study, we provide experimental evidence on the effects of gender quotas in tournaments on performance and selection into a subsequent team task. In particular, participants

[2] Of course, many other important factors such as discrimination or differences in preferences for career choice contribute to the observed discrepancies, too.

in our experiment work on a real-effort calculation task under three different incentive schemes - *piece-rate*, *tournament*, and *team-incentives* – and, in a fourth stage, select whether they want to work under a piece-rate or a team-incentives scheme. Our main experimental manipulation is whether in the tournament winners are solely determined by performance or also by a gender quota that prescribes that in a group of two men and two women at least one women has to be one of the two winners of the tournament. In three different team environments, in which we vary how payoffs are distributed within the team and whether opportunistic behavior in form of effort concealment is possible or not, we then investigate whether having been exposed to the affirmative action policy in the tournament affects performance within teams. Lastly, we test whether the gender quota affects the willingness to work in teams by looking at the incentive scheme choices in the last stage.

We find that, under neither condition, the quota has any detrimental spillover effects on performance or the willingness to work in teams. That is, we find the same level of effort provision, the same amount of effort concealment, and the same fraction of selection into teams irrespective of whether the quota was implemented or not. Furthermore, these effects are very similar for both men and women, i.e., for those who potentially suffer or benefit from the policy. Besides that, and in line with recent evidence by Kuhn and Villeval (2015), we find that women are more attracted to cooperative environments than men. We further demonstrate, however, that the observed gender differences might be subject to subtle features of the team environment as women's higher degree of selection into teams completely disappears when adverse behavior in form of effort concealment is possible.

To the best of our knowledge, this is the first paper that systematically investigates the effects of gender quotas on team performance and the willingness to work in teams.[3] Our paper contributes to the growing literature that evaluates the overall efficacy of affirmative action policies. In particular, it extends previous evidence that finds that such policies do not harm efficiency in tournaments by showing that they do not entail detrimental spillover effects on subsequent team tasks and people's willingness to collaborate with others. This is an important

[3] Two previous papers by Balafoutas and Sutter (2012) and Balafoutas et al. (2016) provide some evidence that point into the direction of no detrimental effects of quotas on effort provision in a team task. The focus of these papers, however, is very different. They also do not study selection into teams and do not consider different team environments.

insight, as very often, e.g., after firm internal promotion tournaments, further cooperation by previous competitors is needed. Our study is related to two previous studies by Leibbrandt et al. (2015) and Fallucchi and Quercia (2016) who investigate the effects of affirmative action on sabotage and retaliation, respectively. However, while both of these studies focus on direct effects of affirmative action on adverse behavior in tournaments, the current study focuses on indirect effects of affirmative action in subsequent tasks.

2. The Experiment

The general structure of our experimental design follows that of previous studies on this topic (compare e.g., Niederle and Vesterlund, 2007; Balafoutas and Sutter, 2012; Niederle et al., 2013) and consisted of four stages. In each stage, subjects had to work on a real-effort task, in which they had to add as many sets of five two-digit numbers as possible within 3 minutes (see Figure 1 for a screenshot).[4] The numbers were randomly drawn, and once the participant submitted an answer, a new problem appeared jointly with information on whether the former answer was correct or not. Participants were not allowed to use a calculator, but scratch paper was provided.

Figure 1: The experimental task

At the beginning of the experiment, participants were informed that, at the end of the experiment, one stage would be randomly selected to determine their earnings (plus a show-up fee of €4). The exact content of each stage, however, was only revealed once the previous stage had been completed. In each stage, subjects received written instructions that were additionally read aloud by the experimenter (see Appendix B for an English translation of the instructions). Subjects then had time to ask questions that were answered in private. The design of the four stages was as follows.

[4] Before the start of the first stage, there was a 1 minute practice period in which subjects were given the opportunity to familiarize themselves with the task. This stage was not incentivized.

2.1 Stage 1 – Piece-rate

In Stage 1, participants worked on the calculation task under piece-rate incentives. If chosen for payment, subjects received a piece-rate of €0.50 for each correctly solved calculation. Stage 1 was common to all treatments and performance in this stage will be used as an individual measure for subjects' ability to perform a simple mathematical task.

2.2 Stage 2 – Tournament

In Stage 2, subjects were randomly assigned into groups of four consisting of two men and two women. The gender composition of groups was common knowledge but subjects did not know who was in their group. Group members then had to compete against each other in a tournament. We implemented two types of tournaments in a between-subjects design: *Standard* and *Quota*. In the *Standard* tournament, the winners of the tournament were the two group members with the highest number of correct calculations in Stage 2. In the *Quota* tournament, the best–performing woman was always one of the winners, while the other winner was the best performing subject among the remaining three group members, i.e., there was at least one women and at most one men among the two winners. In both types of tournaments, ties were broken randomly. If Stage 2 was chosen for payment, the two winners of the tournament received €1.00 per correctly solved calculation, while the other two group members did not receive any payment. At the end of Stage 2 subjects were informed about their own performance and whether they won or not, but received no information about the performance of the other players. In the *Quota* tournament, we further decided not to provide any information about whether the quota altered the results or not, as in most of the real-world applications this type of information is rarely available, too.

2.3 Stage 3 – Team incentives

In Stage 3, subjects had to work on the calculation task under team incentives, i.e., they were paid based on the overall performance of the group. Group composition was the same as in Stage 2 (and this was common knowledge). We implemented three treatments in a between-subjects design, aimed at measuring team performance under different environments. In our *Base* treatment, each correctly solved calculation by any group member generated €0.50 for the entire group. At the end of Stage 3, total team output was split equally among all four group members, yielding an individual payoff of €0.125 per correctly solved question. Performance in *Base* serves

as our benchmark treatment testing whether quotas affect effort provision under team incentives in a subsequent real-effort task.

In the *Unequal* treatment, procedures were the same as in *Base* except that benefits were split unequally among group members. In particular, the two winners of the tournament now each received €0.175 per correctly solved question, while the two losers of the tournament received only €0.075 per question that any group member solved correctly. As a result, losers and winners ended up earning different amounts, leading to payoff inequality within the group. This treatment reflects, for example, a situation in which winning the tournament constitutes an internal promotion to a higher rank (e.g. to the role of a team leader) which, in turn, also leads to higher remuneration. Compared to *Base*, we expect the payoff asymmetry in *Unequal* to increase the scope for spillover effects of the quota, as it now potentially has direct effects on the remuneration in the team task, and hence might increase the perceived unfairness of such policy.

In the *Hide* treatment, procedures were the same as in *Base* except that before earnings were split, subjects had the opportunity to conceal their true performance by hiding some (or all) of their correctly solved calculations from the group. Each hidden calculation yielded a private benefit of €0.40 but no benefit to any of the other group members. Note that this treatment resembles a public goods game (with earned endowments) in which hiding effort is individually beneficial, while devoting the whole performance to the team is socially efficient. The decision to hide effort, thus, can be interpreted as shirking or free riding within teams (e.g., working on a different, team-unrelated task). Compared to the other two team-environments, the possibility to hide effort provides subjects with an additional channel through which they can affect team performance. In particular, in *Base* and *Unequal* we can measure the effects of the quota on team performance only via a potential adjustment in subjects' decision to provide effort in the real-effort task. In *Hide*, in contrast, subjects can additionally affect team performance by their decision of how many questions to hide from the group. The concealment decision, thus, serves as an additional test of whether quotas negatively affect team performance. At the end of Stage 3, subjects were informed about their own performance but not about the performance of their group members.

2.4 Stage 4 – Choice

In Stage 4, subjects could determine themselves which payment scheme should be applied to their Stage 4 performance. They could choose between *piece-rate* (as in Stage 1) and *team incentives* (as in Stage 3). Because a participant's choice can be interpreted as her revealed preference for teamwork, Stage 4 allows us to investigate the effects of the gender quota on the willingness to work in teams. To ensure that groups could always be formed, one choice was randomly selected in each group and implemented for all group members, i.e., each group member's choice had a chance of 25% to be implemented.[5] Subjects were then asked to perform the calculation task under the selected payment scheme. If Stage 4 was selected for payment, then participants were paid according to the implemented payment scheme. At the end of Stage 4, participants were told which Stage was payoff relevant and how much they earned in that stage.[6]

Table 1: Overview of the experimental design

<table>
<tr><th>Treatment</th><th>Stage 1:
piece-rate</th><th>Stage 2:
tournament</th><th>Stage 3:
team-incentives</th><th>Stage 4:
choice</th></tr>
<tr><td>Base (n=64)</td><td rowspan="6">Piece-rate</td><td rowspan="3">Standard</td><td>Baseline</td><td rowspan="6">Choice between piece-rate or team incentives for Stage 4 performance</td></tr>
<tr><td>Unequal (n=64)</td><td>Unequal benefits</td></tr>
<tr><td>Hide (n=64)</td><td>With concealment</td></tr>
<tr><td>Base-Quota (n=60)</td><td rowspan="3">Quota</td><td>Baseline</td></tr>
<tr><td>Unequal-Quota (n=64)</td><td>Unequal benefits</td></tr>
<tr><td>Hide-Quota (n=64)</td><td>With concealment</td></tr>
</table>

Our experiment was implemented in a 2x3 between-subjects design, in which we vary the type of tournament (*Standard* or *Quota*) and the type of team incentives (*Base*, *Unequal*, or *Hide*)

[5] This ensures that the choice of each group member was incentive-compatible. Group assignment was the same as in Stage 2 and 3, and this was common knowledge.

[6] In stages 1-3, we further elicited subjects' beliefs about their relative performance within their group. Belief elicitation was not preannounced, but subjects could earn an additional €0.50 if their guess about the own ranking within the group was correct.

participants faced in Stage 2 and Stage 3, respectively. This leads to six experimental treatments: *Base*, *Base-Quota*, *Unequal*, *Unequal-Quota*, *Hide*, and *Hide-Quota*, which are summarized in Table 1. The experiment was conducted using z-Tree (Fischbacher, 2007) and subjects from various disciplines were recruited using ORSEE (Greiner, 2015). We conducted two sessions per treatment with either 28 or 32 subjects, leading to $n = 380$ participants in total.

3. Results

We start our analysis by analyzing task performance in Stage 1 and 2. Similar to previous studies using the same task (e.g., Balafoutas and Sutter, 2012; Niederle et al., 2013; Leibbrandt et al., 2015), we find that men perform slightly better than women. In Stage 1, men solved on average 6.3 questions correctly, while women answered 5.6 questions correctly (see also Table A1 in the Appendix for an overview of the descriptive statistics). This difference, however, does not reach statistically significance using a nonparametric Mann-Whitney test ($p = 0.114$). In Stage 2, men and women significantly increase their performance by 0.8 and 0.7 to 7.1 and 6.4 correctly solved questions, respectively (Sign-rank test, both $p < 0.001$). This increase can be attributed to either learning or the competitive incentives. Besides that, we find no detrimental effects of the quota on overall tournament performance. In the standard tournament subjects answer on average 6.8 questions correctly, which is only slightly and insignificantly higher than 6.6 under the quota tournament (Mann-Whitney test, $p = 0.440$; Table A2). This effect is very similar for men (7.2 vs. 7.0, Mann-Whitney test, $p = 0.887$) and women (6.5 vs. 6.2, Mann-Whitney test, $p = 0.261$), i.e., there is no interaction between gender and the quota with respect to tournament effort (see also Table A2). In the following, we turn to our main research question by investigating how the gender quota affects team performance and the willingness to work in teams.

3.1 The effects of quotas on effort provision in teams

Figure 2 displays the mean performance in Stage 3 for all treatments. Overall, we find no negative spillover effects of the quota on team performance. That is, under all three types of team environments, the number of correctly solved questions is very similar across Standard and Quota tournament. Support for this result comes from regression analysis using the number of correctly solved questions in Stage 3 as dependent variable. As independent variables, we use dummy

variables (and their interaction) for our six treatments. *Quota* is a dummy variable equal to one if the quota was implemented in the tournament and zero otherwise. *Unequal* is a dummy variable equal to one if in the team task in Stage 3 benefits were distributed unequally, while *Hide* is a dummy variable equal to one if in the team task in Stage 3 hiding of effort was possible and zero otherwise. *Quota x Unequal* and *Quota x Hide* are interaction terms between the type of tournament and the type of team environment. We further control for subjects' ability, measured as the number of correctly solved questions in Stage 1. The results of this regression are reported in Table 2 model (1).

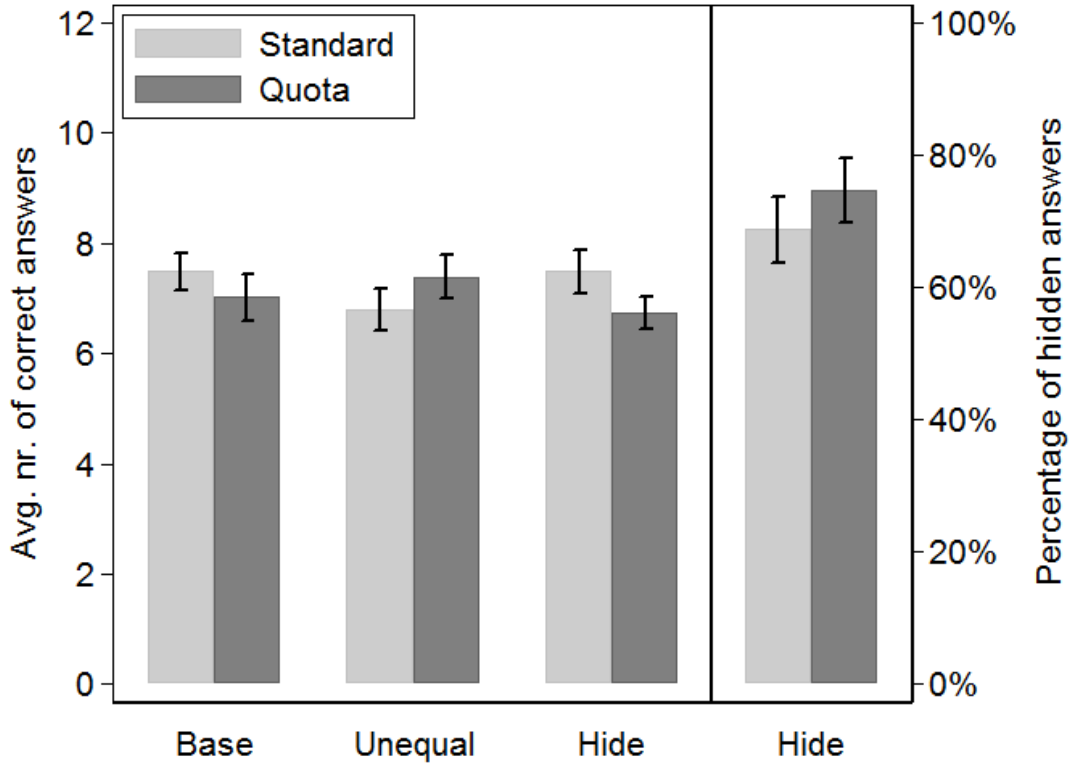

Figure 2: Average number of correctly solved calculations (left) and average percentage of hidden answers (right) in Stage 3 by treatment. Bars indicate standard errors.

In neither treatment, we find the quota to affect team performance as indicated by the insignificant Quota dummy as well as the insignificant interaction terms. That is, the number of correctly solved questions in the team task in Stage 3 is very similar across both types of preceded tournaments. Team performance is also very similar across all types of team environments,

although effort in *Unequal* is somewhat lower compared to *Base* as indicated by the weakly significant Unequal dummy.

In models (2) and (3), we disentangle the effect of the quota by gender, as it is conceivable that men, who are the potential victims of the quota, react differently to the implementation of such policy than women, who potentially benefit from it. Yet, we find the effect of the quota to be very similar across genders and to be statistically insignificant in both cases. Subjects' ability, in contrast, has the expected positive sign and is highly significant in all models. We summarize these findings in our first result.

RESULT 1: *Affirmative action policies in form of gender quotas have no negative spillover effects on subsequent team performance, neither for men nor for women.*

Table 2: Determinants of team performance in Stage 3

Dependent variable:	All	Men	Women
# solved questions in Stage 3	(1)	(2)	(3)
Unequal	-0.594*	-0.818*	-0.380
1 if Unequal, 0 otherwise	(0.318)	(0.475)	(0.441)
Hide	-0.163	-0.086	-0.246
1 if Hide, 0 otherwise	(0.343)	(0.514)	(0.445)
Quota	-0.318	-0.154	-0.492
1 if Quota, 0 otherwise	(0.380)	(0.575)	(0.495)
Quota × Unequal	0.586	0.585	0.602
	(0.521)	(0.786)	(0.696)
Quota × Hide	0.010	-0.453	0.482
	(0.520)	(0.769)	(0.690)
Productivity	0.745***	0.736***	0.746***
# correct questions Stage 1	(0.036)	(0.048)	(0.059)
Constant	3.028***	3.259***	2.846***
	(0.287)	(0.458)	(0.381)
Observations	380	190	190
R^2	0.540	0.574	0.482

Notes: OLS regressions. The dependent variable is the number of correctly solved questions in the team task in Stage 3. Productivity is measured as the number of correct solved calculations under piece-rate incentives in Stage 1. Robust standard errors are in parentheses. Significance levels * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

3.2 The effects of quotas on effort concealment

Our results so far have demonstrated that the implementation of the gender quota does not negatively affect effort provision in our real-effort calculation task. As a next step, we investigate whether the quota has a negative effect on effort concealment in our *Hide* treatments. This is interesting as it gives subjects an additional opportunity of shirking that is different from the decision to withhold effort in our calculation task as in the other treatments. The results of this treatment can hence serve as an additional test of whether quotas negatively affect team performance.

Table 3: Determinants of effort concealment

Dependent variable:	% hidden effort		
	(1)	(2)	(3)
Quota *1 if Quota, 0 otherwise*	5.958 (6.976)	3.603 (10.481)	7.388 (6.174)
Female *1 if Female, 0 otherwise*		1.086 (9.940)	0.978 (4.372)
Female × Quota		4.710 (14.044)	-5.998 (8.092)
Productivity *# correct questions Stage 1*			0.840 (0.722)
Belief Hide			1.022*** (0.084)
Constant	68.711*** (4.931)	68.168*** (7.471)	-10.186 (9.253)
Observations	128	128	128
R^2	0.006	0.009	0.650

Notes: OLS regressions. The dependent variable is the percentage number of hidden questions from the group in Stage 3. Productivity is measured as the number of correct solved calculations under piece-rate incentives in Stage 1. Belief Hide is a subject's belief about the average percentage of hidden answers by the group members, measured in steps of 10 percentage points. Robust standard errors are in parentheses. Significance levels * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Overall, we find a remarkably high level of 72% hidden answers with only small differences between standard and quota tournaments (69% and 76%, respectively).[7] To test the significance of this result, we run OLS regressions with the percentage of hidden questions as dependent variable and a dummy for the Quota tournament as independent variable. We report the results of this regression in model (1) in Table 3. While it confirms the somewhat higher hiding rates after quota than after standard tournaments, this difference is not statistically significant as indicated by the insignificant Quota dummy. Hence, even when providing subjects with an additional channel through which they can alter team performance, we find no negative spillover effects of the quota.

In model (2), we, again, disentangle this effect by gender to test whether those who benefit and those who suffer from the affirmative action policy react differently. We find no evidence for such an effect. Overall, men and women hide 70% and 73% of their answers, respectively, and these differences are not statistically significant, neither after standard nor after quota tournaments. Finally, model (3) shows that hiding of effort does not depend on subject's productivity, but that subjects act on average reciprocally, i.e., they hide the more answers the higher their belief about their group members' hiding decision.[8] We summarize these findings in our second result.

> **RESULT 2:** *The quota has no negative effect spillover effect on effort concealment. Men and women hide about the same level of their effort.*

3.3 The effects of quotas and team environment on selection into teams

Figure 3 summarizes the choice decision over the preferred incentive scheme in Stage 4. Panel A displays the percentage of subjects choosing team incentives for all six treatments. It shows that the quota has no systematic effect on selection into teams. While in *Base* and *Unequal* the quota slightly decreases selection into teams, in *Hide* it slightly increases the percentage of subjects choosing team incentives. These effects, however, are relatively small in magnitude and

[7] These results suggest a somewhat lower cooperation level than typically observed in one-shot public goods games where it is about 50%. One possible reason for this result is that having been exposed to a competitive environment immediately before making a cooperation decision decreases cooperation as shown by Buser and Dreber (2016).

[8] Beliefs were elicited by asking subjects to state the percentage of answers they think the other group members hide, on average. Percentages had to be stated in steps of 10 percentage points, and subjects were rewarded a bonus of €0.50 if their guess was correct. Beliefs were not preannounced and were elicited only after subjects' hiding decision.

not statistically significant (*Base*: $\chi^2(1) = 0.67$, $p = 0.412$; *Unequal*: $\chi^2(1) = 0.54$, $p = 0.461$; *Hide*: $\chi^2(1) = 0.67$, $p = 0.414$). Overall, the percentage of subjects willing to work in teams after standard and quota tournaments is almost identical and equal to 34% and 32%, respectively ($\chi^2(1) = 0.26$, $p = 0.611$). These findings constitute our third result.

RESULT 3: *Gender quotas do not affect selection into teams.*

The type of team environment in contrast, significantly affects the attractiveness of teamwork ($\chi^2(2) = 6.01$, $p = 0.049$). In particular, while the fraction of subjects choosing team-incentives is very similar across *Base* and *Unequal* (39% vs. 36%, $\chi^2(1) = 0.21$, $p = 0.649$), the possibility to behave opportunistically by hiding effort significantly decreases selection into teams to 25% ($\chi^2(1) = 5.43$, $p = 0.019$ and $\chi^2(1) = 3.61$, $p = 0.057$ compared to *Base* and *Unequal*, respectively).

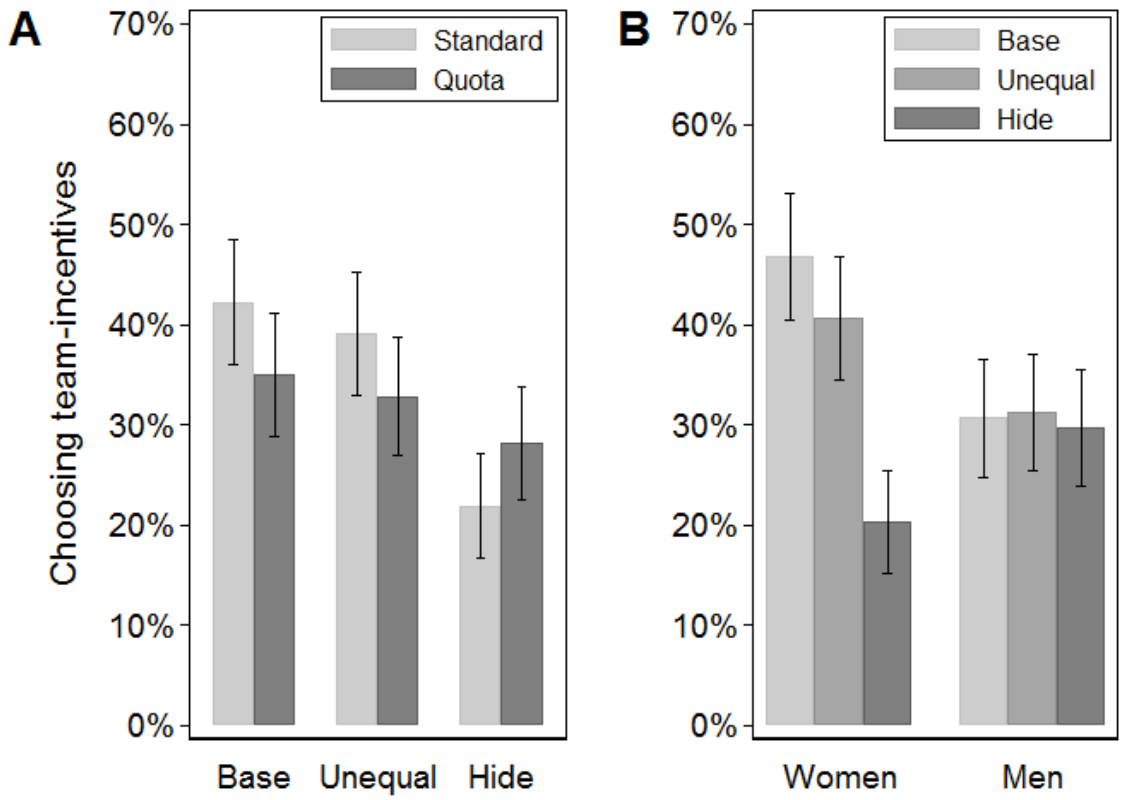

Figure 3: Fraction of subjects selecting team incentives in Stage 4 by treatment (Panel A) and gender (Panel B). Bars indicate standard errors.

As a last step, in Panel B of Figure 2, we compare the choice of the preferred incentive scheme in Stage 4 by gender. Because the previous results have revealed no significant differences across the standard and the quota tournament, in the following, we pool observations across both types of tournaments. While there is ample of evidence that men are more attracted to competitive environments than women (e.g., Niederle and Vesterlund, 2007), much less is known about gender differences in the selection into different environments. One exception is a recent study by Kuhn and Villeval (2015), who find that women are more likely to select into team incentives than men. Building on this finding, here we test how sensitive men and women react to changes in the team environment and whether women's higher willingness to work in teams extends to the different environments such as in *Unequal* and *Hide*.

Strikingly, we find that the observed differences between the different types of team environments are entirely driven by women. While men select into teams equally often in *Base*, *Unequal*, and *Hide* (31%, 31%, and 30%, respectively; $\chi^2(2) = 0.04$, $p = 0.981$), women's willingness to work in teams significantly reduces by more than half when the opportunity of hiding effort is available (47%, 41%, and 20%, $\chi^2(2) = 10.28$, $p = 0.005$). As a result, while under standard team-incentives females are more likely to select into team than men (47% vs. 31% $\chi^2(1) = 3.40$, $p = 0.065$), this effect vanishes and even turns around when hiding is possible, although the difference does not reach statistical significance (20% vs. 30%, $\chi^2(1) = 0.15$, $p = 0.221$).[9] These results are further confirmed by logistic regression analysis, which we report in Table A3 in the Appendix. We summarize our findings as follows.

> **RESULT 4:** *Irrespective of the team environment, about one third of all men prefer to work under team-incentives rather than under a piece-rate. Women, in contrast, condition their willingness to work in teams on the specific nature of the team environment; they significantly less often select into teams when opportunistic behavior in form of hiding effort is possible.*

[9] This result is not driven by different expectations about the hiding decision of others as men and women have very similar beliefs about the average percentage of hidden answers by their group members (men: 65%, women: 70%, Mann-Whitney test, $p = 0.516$)

Hence, in line with the recent evidence by Kuhn and Villeval (2015), we find that women are more attracted to cooperative environments than men, but further show that the observed gender differences might be subject to subtle features of the work environment. In particular, women's larger attraction to teams completely disappears when adverse behavior is possible.

4. Discussion and Conclusion

Affirmative action policies such as quotas have been proposed and implemented as one tool to reduce the gender imbalance in labor market outcomes. Previous experimental work has examined the effects of such programs on tournament performance, finding no negative effects (Balafoutas and Sutter, 2012; Niederle et al., 2013; Calsamiglia et al., 2013). Much less is known, however, about potential detrimental effects of affirmative action policies in other environments, e.g., due to perceived unfairness.

Building on recent evidence (Balafoutas and Sutter, 2012; Balafoutas et al., 2016), here we experimentally test whether such policies in form of gender quotas have negative spillover effects on subsequent performance within teams, as well as on the willingness to work in teams. Comparing three different types of team environment, we find no detrimental effects of the quota, neither on team performance nor on selection into teams. While this is good news for advocates of such policies, more research is needed to evaluate the overall efficacy of affirmative action programs. A recent study by Leibbrandt et al. (2015), for example, shows that gender quotas can lead to a backlash against women by increased sabotage in tournaments, and Fallucchi and Quercia (2016) find that affirmative action policies are less effective in promoting disadvantaged groups when people have the opportunity to retaliate. Future research should therefore test the robustness of these results. For example, it would be interesting to see whether different types of affirmative action policies that are more (or less) justifiable than gender quotas or using a task that is not gender neutral lead to different results. Furthermore, in order to formulate concrete policy recommendations, it is important to test whether our results generalize in more natural field settings, in which teamwork comprises much more than the anonymous one-shot interaction, as studied in our experiment.

Bibliography

Balafoutas, L., Davis, B. J., & Sutter, M. (2016). Affirmative action or just discrimination? A study on the endogenous emergence of quotas. *Journal of Economic Behavior & Organization*, *127*, 87-98.

Balafoutas, L., & Sutter, M. (2012). Affirmative action policies promote women and do not harm efficiency in the laboratory. *Science*, *335*(6068), 579-582.

Blau, F., & Kahn, L. (2000). Gender Differences in Pay. *Journal of Economic Perspectives*, *14*(4), 75-99.

Blau, F. D., Ferber, M. A., & Winkler, A. E. (2013). *The economics of women, men and work*. Pearson Higher Ed.

Booth, A., & Nolen, P. (2012). Choosing to compete: How different are girls and boys?. *Journal of Economic Behavior & Organization*, *81*(2), 542-555.

Brandts, J., Groenert, V., & Rott, C. (2014). The impact of advice on women's and men's selection into competition. *Management Science*, *61*(5), 1018-1035.

Buser, T., & Dreber, A. (2016). The flipside of comparative payment schemes. *Management Science*, 62(9), 2626-2638

Buser, T., Niederle, M., & Oosterbeek, H. (2014). Gender, competitiveness and career choices. *The Quarterly Journal of Economics*, *129*(3), 1409-1447.

Calsamiglia, C., Franke, J., & Rey-Biel, P. (2013). The incentive effects of affirmative action in a real-effort tournament. *Journal of Public Economics*, *98*, 15-31.

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic literature*, *47*(2), 448-474.

Datta Gupta, N., Poulsen, A., & Villeval, M. C. (2013). Gender matching and competitiveness: Experimental evidence. *Economic Inquiry*, *51*(1), 816-835.

Fallucchi, F., & Quercia, S. (2016). Affirmative action and retaliation in experimental contests. *CBESS Discussion Paper No 16-03*.

Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics*, *118*(3), 1049-1074.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association*, *1*(1), 114-125.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171-178.

Hamilton, B. H., Nickerson, J. A., & Owan, H. (2003). Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of Political Economy*, *111*(3), 465-497.

Kuhn, P., & Villeval, M. C. (2015). Are women more attracted to cooperation than men?. *The Economic Journal*, *125*(582), 115-140.

Lazear, E. P., & Shaw, K. L. (2007). Personnel economics: The economist's view of human resources. *The Journal of Economic Perspectives*, *21*(4), 91-114.

Leibbrandt, A., Wang, L. C., & Foo, C. (2015). Gender quotas, competitions, and peer review: experimental evidence on the backlash against women. *CESifo Working Paper Series No. 5526*.

Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much?. *The Quarterly Journal of Economics*, 1067-1101.

Niederle, M., Segal, C., & Vesterlund, L. (2013). How costly is diversity? Affirmative action in light of gender differences in competitiveness. *Management Science*, *59*(1), 1-16.

Schotter, A., & Weigelt, K. (1992). Asymmetric tournaments, equal opportunity laws, and affirmative action: Some experimental results. *The Quarterly Journal of Economics*, 511-539.

Sutter, M., & Glätzle-Rützler, D. (2015). Gender differences in the willingness to compete emerge early in life and persist. *Management Science*, *61*(10), 2339-2354.

Villeval, M. C. (2012). Ready, steady, compete. *Science*, *335*(6068), 544-545.

Weichselbaumer, D., & Winter-Ebmer, R. (2007). The effects of competition and equal treatment laws on gender wage differentials. *Economic Policy*, *22*(50), 236-287.

Appendix

A. Additional tables

Table A1: Mean performance in Stages 1-3 by treatment and gender.

Treatment	**Stage 1** *piece-rate*			**Stage 2** *tournament*			**Stage 3** *team incentives*		
	Men	Women	Total	Men	Women	Total	Men	Women	Total
Base (n=64)	6.91 (0.55)	5.06 (0.36)	5.98 (0.35)	7.56 (0.57)	6.59 (0.37)	7.08 (0.34)	8.34 (0.50)	6.63 (0.37)	7.48 (0.33)
Base-Quota (n=60)	5.83 (0.35)	5.73 (0.64)	5.78 (0.36)	7.33 (0.52)	6.10 (0.61)	6.72 (0.41)	7.40 (0.51)	6.63 (0.69)	7.02 (0.43)
Unequal (n=64)	5.94 (0.69)	5.78 (0.48)	5.86 (0.42)	6.69 (0.69)	6.16 (0.42)	6.42 (0.40)	6.81 (0.70)	6.78 (0.37)	6.80 (0.39)
Unequal-Quota (n=64)	6.63 (0.71)	5.97 (0.36)	6.30 (0.40)	7.03 (0.65)	6.25 (0.48)	6.64 (0.41)	7.75 (0.65)	7.03 (0.47)	7.39 (0.40)
Hide (n=64)	6.34 (0.62)	6.06 (0.35)	6.20 (0.35)	7.28 (0.62)	6.75 (0.47)	7.02 (0.39)	7.84 (0.63)	7.13 (0.49)	7.48 (0.40)
Hide-Quota (n=64)	5.94 (0.40)	5.28 (0.42)	5.61 (0.33)	6.66 (0.40)	6.31 (0.47)	6.48 (0.34)	6.94 (0.43)	6.53 (0.41)	6.73 (0.30)

Notes: Performance is measured as the number of correctly solved calculations. Standard errors are in parentheses.

Table A2: Effects of quota on effort provision in the tournament in Stage 2.

	(1)	(2)
Quota *1 if Quota, 0 otherwise*	-0.138 (0.209)	-0.295 (0.275)
Female *1 if Female, 0 otherwise*		-0.099 (0.290)
Quota x Female		-0.314 (0.419)
Productivity	0.764*** (0.039)	0.760*** (0.039)
Constant	2.241*** (0.266)	2.315*** (0.322)
Observations	380	380
R^2	0.548	0.550

Notes: OLS regressions. The dependent variable is the number of correctly solved calculations in the tournament in Stage 2. Productivity is measured as the number of correct solved calculations under the piece-rate incentives in Stage 1. Robust standard errors clustered at the individual level are in parentheses. Significance levels * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Determinants of selection into teams in Stage 4

	(1)	(2)
Unequal *1 if Unequal, 0 otherwise*	-0.130 (0.361)	0.028 (0.386)
Hide *1 if Hide, 0 otherwise*	-0.958** (0.395)	-0.045 (0.389)
Quota *1 if Quota, 0 otherwise*	-0.304 (0.371)	
Quota x Unequal	0.032 (0.524)	
Quota x Hide	0.639 (0.554)	
Female *1 if Female, 0 otherwise*		0.688* (0.376)
Female × Unequal		-0.279 (0.528)
Female × Hide		-1.192** (0.559)
Constant	-0.315 (0.253)	-0.817*** (0.276)
Observations	380	380
Log Likelihood	-237.388	-147.831

Notes: Logistic regressions. The dependent variable takes the value 1 if a subject chose team-incentives, and zero otherwise. Robust standard errors are in parentheses. Significance levels * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B. Experimental instructions

Instructions on paper (translated from German)

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 enter your decisions on the keyboard. All decisions and answers will remain confidential and anonymous. The payments at the end of the experiment will also be anonymous, i.e., no one will receive any information about the payments of others.

The experiment consists of 4 stages. One of these four stages (1-4) will be randomly selected 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 €4.

You will receive instructions for each of the four 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 unclear.

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

Stage 1 – Piece rate *[same in all treatments]*

Your task in stage 1 is to solve correctly as many addition exercises as possible. To be more precise, you will have 3 minutes' 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 *[Standard]*

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 4 participants, 2 of whom are men and 2 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. 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 three 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 **€1.00 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.

At the end of this stage, you will be informed about whether you have won or not.

Stage 2 – Tournament *[Quota]*

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 4 participants, 2 of whom are men and 2 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. 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.**

The two winners of the tournament are determined as follows In each group, one of the two winners is in any case the woman with the best performance (of all two women). The other winner is the group member with the best performance among the remaining members (i.e., excluding the best-performing woman). The two winners receive **€1.00 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.

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

Summary:

A woman wins the tournament 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 the tournament 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. Therefore, there is at least one woman and at most one man as winners in the tournament.

At the end of this stage, you will be informed about whether you have won or not.

Stage 3 – Team *[Base]*

As in stages 1 and 2, you will have again 3 minutes' time in order to solve correctly as many addition exercises as possible. The group composition (with 2 men and 2 women) is the same as in stage 2. As in stage 2, your payment in this stage depends on your performance as well as on the total performance of all other members in your group; however, in a different manner as in stage 2.

If stage 3 is the stage selected for payment (among stages 1-4), then your payment is determined as follows. You receive **12.5 eurocent** 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 4 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 4 group members**.

After stage 4, you will be informed about the performance of the whole group, which will determine your payment from this stage. Your payment is not reduced when you (or any of your group members) enter a wrong answer. From now on, we call this method of payment the Team payment.

Stage 3 – Team *[Unequal]*

As in stages 1 and 2, you will have again 3 minutes' time in order to solve correctly as many addition exercises as possible. The group composition (with 2 men and 2 women) is the same as in stage 2. As in stage 2, your payment in this stage depends on your performance as well as on the total performance of all other members in your group; however, in a different manner as in stage 2.

If stage 3 is the stage selected for payment (among stages 1-4), then your payment is determined as follows. The two **winners from stage** 2 receive **17.5 eurocent** for each correct answer that a member of your group has entered in the 3 minutes. The two **losers from stage 2** receive **7.5 eurocent** 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 4 members). It also means that the two winners will receive a higher payment than the two losers and that the two winners will receive the same payment and that the two losers will receive the same payment. The exact payment depends thereby **only on the total performance of the group, i.e., on the sum of all correct answers of the 4 group members**.

After stage 4, you will be informed about the performance of the whole group, which will determine your payment from this stage. Your payment is not reduced when you (or any of your group members) enter a wrong answer. From now on, we call this method of payment the Team payment.

Stage 3 – Team *[Hide]*

As in stages 1 and 2, you will have again 3 minutes' time in order to solve correctly as many addition exercises as possible. The group composition (with 2 men and 2 women) is the same as in stage 2. As in stage 2, your payment in this stage depends on your performance as well as on the total performance of all other members in your group; however, in a different manner as in stage 2.

If stage 3 is the stage selected for payment (among stages 1-4), then your payment is determined as follows. At the end of stage 3, you will be informed about how many additions you solved correctly. You then have to decide how many of these additions you want to **hide** from the group. For each answer you hide, you receive **40 eurocent** and the other group members receive **0 eurocent.** Conversely, you also receive no payment from answers your group members hide. For each answer you **do not hide**, you and each of the other three group members receive **12.5 eurocent**. Likewise, you and each of the other group members receive **12.5 eurocent** for each correct answer one of your group members does not hide. That means that each answer that is **not hidden** is worth **€ 0.50** for the **group as a whole** (i.e. all 4 group members together).

After stage 4, you will be informed about the performance of the whole group, which will determine your payment from this stage. Your payment is not reduced when you (or any of your group members) enter a wrong answer. From now on, we call this method of payment the Team payment.

Stage 4 – Choice *[same in all treatments]*

As in stages 1, 2 and 3, you will have again 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 4. You can choose either the **Piece-rate payment** (as in stage 1) or the **Team payment** (as in stage 3). Your group members will make the same decision between Piece-rate payment and the Team payment. The group composition (with 2 men and 2 women) is the same as in stages 2 and 3. After all group members have made their decision, one group member will be randomly selected and his decision will be implemented. That means that the decision of this group member determines the payment method (Piece-rate payment or the Team payment) that is implemented for the entire group in stage 4.

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

- If the **Piece-rate payment** is the implemented payment method, then you will receive **€0.50 for each correct answer.**
- If the **Team payment** is the implemented payment method, then the exact same rules as in stage 3 apply. Therefore, please have another look at the instructions of stage 3.

You will first be asked whether you want to choose the Piece-rate payment or the Team payment for your performance in stage 4. After that, you will be informed about which payment method will be implemented and count in stage 4. Then, you will have 3 minutes' time to solve correctly additions of two-digit numbers. At the end, you will receive information about your payment from stage 4.