DICE

DISCUSSION PAPER

NO 405

Who’s Who: How Uncertainty About the Favored Group Affects Outcomes of Affirmative Action

Chi Trieu

August 2023

hhu
Heinrich Heine
Universität Düsseldorf

IMPRINT

DICE DISCUSSION PAPER

Published by:
Heinrich-Heine-University Düsseldorf,
Düsseldorf Institute for Competition Economics (DICE),
Universitätsstraße 1, 40225 Düsseldorf, Germany
www.dice.hhu.de

Editor:
Prof. Dr. Hans-Theo Normann
Düsseldorf Institute for Competition Economics (DICE)
Tel +49 (0) 211-81-15125, E-Mail normann@dice.hhu.de

ISSN 2190-9938 (online) / ISBN 978-3-86304-404-6

Who's Who: How Uncertainty About the Favored Group Affects Outcomes of Affirmative Action*

Chi Trieu†

August 2023

Abstract

When affirmative action policies target more than one disadvantaged group, they contain uncertainty as to whether an individual who belongs to one of these groups was actually favored. In a laboratory experiment, we study how this feature affects outcomes of affirmative action in the form of quotas, and compare it with two other conditions, namely affirmative action with a certain favored group and no affirmative action. We find that when a group is favored with certainty and the social identity that triggers affirmative action is made salient, affirmed individuals are wrongly perceived as less competent, both by themselves and by others. Consequently, their willingness to compete does not increase and they are selected less for teamwork post competition. Affirmative action with uncertain favored groups does not distort belief in competence, and thus does not induce such unintended consequences. In contrast, it increases competition entry of the affirmed groups and enhances their chances of being selected for teamwork.

Keywords: Affirmative action, Competition, Uncertainty, Identity, Experiment.
JEL Classification: C91, D02, J71

*We thank Marta Serra Garcia, Marc Möller, Lionel Page, Hannah Schildberg-Hörisch, Jana Willrodt, participants at ESA World Meeting 2020, and two anonymous reviewers for valuable comments and suggestions. Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through 235577387/GRK1974 is gratefully acknowledged.

†Düsseldorf Institute for Competition Economics (DICE) and Compass Lexecon, Kö-Bogen, Königsallee 2b, 40212 Düsseldorf, Germany. Email address: ctrieu@compasslexecon.com.

1 Introduction

Affirmative action is implemented by institutions around the globe.[1] Nevertheless, it remains a highly controversial policy, even among its proponents. One critical opinion points to the potential over- and under-inclusiveness of conventionally practiced affirmative action policies targeting one specific disadvantaged group. Notably, favoring one particular minority seems subjective and might fail to fully capture the truly disadvantaged individuals in an increasingly diverse population and workforce.

This paper aims to experimentally investigate one key feature of the simultaneous occurrence of several affirmative action policies, namely the uncertainty about the targeted groups. Clearly, firms, governments, and universities as decision-makers know how affirmative action policies influence outcomes of hiring, promotion, or admission decisions. However, this information is usually not publicly communicated. Hence, from the applicants' side, there is an inherent uncertainty whether or not they were actually favored, even if there does exist an affirmative action policy explicitly targeting them. We build on this key feature and examine how it influences multiple outcomes of affirmative action policies in the form of quota rules in a laboratory experiment.

Affirmative action is typically studied in a competitive setting to resemble real-world admission, hiring, and promotion processes. In such settings, the uncertainty and salience of identity play a role in how individuals and groups respond to competitive incentives (see section 2). Despite the emerging empirical literature on the effectiveness and consequences of affirmative action policies, much less is known about the relationship between these factors and outcomes of affirmative action, which is the main focus of this study.

In a competitive setting, we form identities by characterizing each subject by gender and a randomly given color (green or blue). Both gender and the stylized color characteristics allow us to create completely distinct groups, which is a useful feature for our analysis.[2] Our main treatment variation is the occurrence and criteria of affirmative action policies. In the baseline treatment, no quota is implemented and thus no specific social identity is enhanced. In a second treatment, a 50% gender quota is applied to target women, making the identity of the favored group salient and fully observable. In a third treatment, either women or one color group are favored by a 50% quota rule with a fifty-fifty chance. As such, it is uncertain whether women are the favored group, and the identity of the (non)favored

[1] See, e.g., Holzer and Neumark (2000) for affirmative action policies in the U.S., Sowell (2004) for affirmative action policies in India, Malaysia, Nigeria, and Sri Lanka, and Hyman et al. (2012) for diversity management tools, including affirmative action policies under the legislation of the European Union.

[2] It is noteworthy that artificially inducing identity does not reduce the degree of discrimination in lab experiments compared to using natural identities (see, e.g., Lane (2016) for a meta-analysis).

group is not salient. Our assignment of group identities is crucial for causal identification, yet it is surely stylized compared to the competition in the labor market or admission to higher education. In section 5, we discuss the implications as well as the possible limitations of our approach.

We examine both immediate and spillover effects under different policies. First, we test the immediate outcomes by comparing efficiency in terms of task performance and how candidates are selected, as well as willingness to compete in tournaments across treatments. Second, we investigate spillover outcomes in a following teamwork setting. While building on the classical design of Niederle and Vesterlund (2007), our experiment contributes in two ways: (i) we introduce uncertainty about the favored group, thus varying the salience of the social identity of the beneficiaries across treatments, and (ii) we add perceived competence as a new measure of the spillover outcomes of affirmative action.

We find that uncertainty about the favored group does not affect efficiency measured by performance. When affirmative action explicitly targets only gender, it does not encourage women to enter competition. In contrast, when both women and one color group are targeted, competition entry of the favored groups significantly increases. Controlling for risk preference, we propose that this effect operates through beliefs. Women under gender quotas are significantly less confident about their relative rank compared to women and favored types in other treatments. The difference in belief is not attributed to an actual difference in performance, suggesting that gender quotas lead women to have less self-confidence in their competence, thus lowering their willingness to compete. Interestingly, there is evidence of a similar effect for men under gender quotas.

In the post-competition setting, we find a general misperception of competence toward women in the baseline treatment. They are selected significantly less for teamwork, although their performance is not lower than men's. More strikingly, this gap becomes wider under gender quotas, but narrower and insignificant when women are not favored with certainty.

Both immediate and post-competition outcomes of different quotas lead us to conclude that the salience of identity of the favored group plays a role in determining the consequences of affirmative action policies. When there is one favored group and its identity is easy to make salient, affirmed individuals might be perceived as less competent, both by themselves and by others. This is not the case when the identity of the favored groups is uncertain. Our findings imply that negative consequences of gender quotas observed in previous studies with women as the solely favored group (Leibbrandt, Wang, and Foo (2017); Ip, Leibbrandt, and Vecci (2020)) might well describe the "worst-case scenario".

Furthermore, simultaneously targeting several disadvantaged groups and avoiding priming group identities as the (non)favored play an important role in making affirmative action work.

The rest of the paper proceeds as follows. Section 2 puts our study into perspective with the related literature, section 3 explains the experimental design, section 4 shows the experimental results, and section 5 discusses the implications of our main results and concludes.

2 Related literature

Our research is directly related to a growing line of empirical literature evaluating consequences of affirmative action policies. A substantial amount of literature along the same line studies quotas for women, racial and ethnic minorities, resembling real-world policies targeting these groups.[3] We contribute by investigating outcomes of affirmative action policies when there are two targeted groups, and the affirmed group is uncertain.[4]

More broadly, our study bridges the affirmative action literature with the literature on identity and uncertainty in competitive settings. A more salient group identity is linked to changes in individual willingness to compete (Gupta, Anders, and Villeval (2013); Cornaglia, Drouvelis, and Masella (2019)), more intense competition between groups (Chen, Ong, and Sheremeta (2015); Kato and Shu (2016)), and increasing group conflicts (Chowdhury, Jeon, and Ramalingam (2016)). Balafoutas and Sutter (2019) show that uncertainty or ambiguity in number of winners strongly shifts the gender differences in performance and willingness to compete in favor of male participants. Gee (2019) finds that simply removing the uncertainty about the number of applicants increases application finish rate, and encourages women and minorities to apply for jobs.

Outcomes of affirmative action policies documented in experimental studies are by far ambiguous.[5] The pros and cons of affirmative action pave the way for a new strand of studies emphasizing the importance of providing evidence of discrimination against the targeted group (Ip, Leibbrandt, and Vecci (2020)), justifications (Petters and Schroeder (2020)), and fairness perceptions (Schildberg-Hörisch et al. (2023)) in guaranteeing positive outcomes of affirmative action policies. Our study supplements research in this vein by adding the perspective of uncertainty and the salience of social identity.

[3]See Appendix A for a more comprehensive review of this literature.
[4]See Appendix A for specific examples.
[5]See Appendix A for further details.

3 Experimental Design

3.1 The real-effort task

We apply affirmative action on the performance of an arithmetic task introduced by Niederle and Vesterlund (2007) in which subjects have to manually add as many sets of five 2-digit numbers as possible in a given time. After each attempt, subjects move directly to the next questions without getting feedback on whether their answer was correct.

Measuring performance by the arithmetic task serves two purposes. First, no special knowledge is required, and a minimal learning effect has been documented in previous studies (Niederle and Vesterlund, 2007; Balafoutas and Sutter, 2012). Second and more importantly, previous studies show that gender difference in willingness to compete is sensitive to the nature of the tasks. In particular, women are less competitive in mathematical tasks than men (see, e.g., Sutter, Zoller, and Glätzle-Rützler (2019)). Thus, using this task to measure performance, we introduce a reasonable need for gender quotas in our setup.

3.2 Stages and treatment design

The experiment consists of five stages and a final questionnaire. Each stage is designed as follows:

Stage 1: Piece-rate. Subjects work on the task for five minutes and receive a piece-rate payment of 0.5 EUR per correct answer. Before starting stage 1, we give subjects a 2-minute trial round to eliminate the learning effect. The purpose of stage 1 is to measure the baseline performance without competitive incentives, as part of the efficiency assessment together with stage 2 and stage 3. As will be explained further in the Result section, we evaluate efficiency by cross-treatment comparison of the average correct tasks of two groups - the candidate pools and the winners of the competitions. Performance is measured in terms of the average number of correctly solved tasks. Each group's performance is observed under two settings, with and without competitive incentives. The competitive incentives are further broken down into compulsory (stage 2) and self-selected (stage 3). Stage 1 performance serves as the measure of performance without competitive incentives..

Stage 2: Tournament. In stage 2, subjects work under tournament incentives. Each subject competes in a group of six, with three men and three women, three Blue types and three Green types. Each subject has a 50% chance of having either color, independent of gender. Subjects stay in the same group throughout the experiment. Two winners are selected in each group. Each winner earns 1.5 EUR per correct answer, while the losers earn nothing. A random tie-breaking rule is applied in the case of a tie. The purpose of this stage is to assess efficiency together with stage 1 and stage 3, and the composition of winners. In terms of efficiency, stage 2 measures the effect of different affirmative action schemes on performance of the candidate pool and the winners under compulsory competition. The composition of winners is shown by the number of each type (gender x color) in the pool of winners.

Subjects are informed about their types at the beginning of stage 1, and group structure is informed at the beginning of stage 2. Subjects never learn the type of other subjects. To avoid any wealth effects, we inform the results of the tournament at the end of the experiment.

We implement a *between-subject treatment design* which determines how the winners are selected.

1. **Baseline treatment (CTR)**: winners are the two best performers.
2. **Gender quotas treatment (GQ)**: there must be at least one woman among the two winners.
3. **Mixed quotas treatment (MIX)**: gender quota is applied with a 50% chance, and color quota is applied with another 50% chance. Under the gender (color) quotas, there must be at least one woman (one Green type) among the two winners.

The treatment design implies that in all treatments, the best performer is always one of the two winners. If the quota rules are not fulfilled, the second-best winner is replaced by the best performer from the favored groups.[6]

Stage 3: Choice between piece-rate and tournament. After subjects have experienced both piece-rate and tournament incentives, we ask them to choose which scheme they want to apply to their score in this stage. If the tournament incentive is chosen, their score is compared to stage 2 score of the other five group members, irrespective of their choices (Niederle and Vesterlund (2007)). Winners are then selected with the same rules and payoffs as in stage 2. If the piece-rate is chosen, subjects receive 0.5 EUR per correct answer. Stage 3 is to elicit the willingness to compete and to assess efficiency together with stage 1 and stage 2. On willingness to compete, stage 3 measures the proportion of favored

[6] We make sure that subjects understand the rule of the competition via several control questions. Each control question is about a made-up competition with information on performance and the favored group. Subjects need to correctly select the winners in all control questions before they can start working on the task.

and non-favored types self-selecting into completions. On efficiency, this stage measures performance under the self-selected competition of the candidate pool and winners.

At the end of stages 1, 2, and 3, subjects receive feedback on their own performance in the respective stage. Subjects never know about the performance and choices of other subjects.

Stage 4: Submit stage 1 score to either piece-rate or tournament. We ask subjects to choose their preferred scheme to apply for their score in stage 1. If the tournament is chosen, their stage 1 score is compared with the stage 1 scores of the other five group members, irrespective of their choices. The winners are then selected with the same rule and payoffs as in stage 2. If the piece-rate is chosen, subjects receive 0.5 EUR per correct answer in stage 1. Asking subjects to submit their past performance to their preferred scheme accounts for preference for performing on top of preference for competing (Ifcher and Zarghamee (2016)). The outcome of stage 4 is independent and does not replace the outcome of stage 1.

Belief elicitation: At the end of stage 4, we elicit beliefs about relative performance in stage 1, stage 2, and stage 3 both within the group of six and within the group of the same gender/color. We incentivise by randomly selecting one guess to be paid. Subjects receive 1 EUR if their selected guess is correct (see Appendix C).

Stage 5: Partner selection under team incentives. Subjects move to a new working setup with a new task (Grid task – Abeler et al. (2011)), while remaining in the same group and the same treatment.[7] In this stage, subjects are first asked to build their own team of three by selecting two among five group members. Each team of three must include one leader and two colleagues. Leaders (colleagues) must be one of two winners (two of four losers) in stage 2. When selecting team members, subjects are informed about the gender of each group member, and whether they are leaders or not. Naming winners (losers) in stage 2 as leaders (colleagues) is to emphasize the hierarchical order from stage 2 tournament outcome. For the same purpose, we introduce unequal bonuses for leaders and non-leaders (Balafoutas, Davis, and Sutter (2016)). Subjects receive a bonus of 5 EUR if they are the leaders, and 2 EUR otherwise.

Subjects then work on the Grid task for five minutes under a team incentive. The overall payoff of the team is equally shared among all members. Each correctly solved grid generates 0.3 EUR for the team, i.e., 0.1 EUR for each member.

[7] In this task, subjects count the number of zeros in a 10-by-10 table containing 100 digits of randomly distributed zeros and ones. Similar to the arithmetic task, this task does not require special prior knowledge or skills.

Subjects do not learn the partner selection of other group members. The choice of a different task allows emphasizing that belief of abilities of other team members is purely perceived rather than based on subjects' experience in the previous stages.[8]

At the end of the experiment, subjects answer a questionnaire which includes our control variables. The questionnaire contains measures on risk and social preferences, positive and negative reciprocity, fairness perception, cognitive ability, overconfidence, and socio-demographics (see Appendix E).

3.3 Procedures

The experiment is registered in the AEA RCT Registry.[9] Overall, we adhered to the registered number of observations as well as the experimental design. We conducted the experiment at the DICE Lab at the University of Düsseldorf in December 2019 and January 2020 using the software zTree (Fischbacher (2007)). Subjects were recruited via ORSEE (Greiner (2004)) from the subject pool of the DICE Lab.

In total, 240 subjects of various disciplines participated in the experiment, with 64, 84, 90 subjects in CTR, GQ, and MIX, respectively. One out of the five stages of the experiment is randomly selected to be payoff-relevant. The total payoff consists of a 4 EUR show-up fee, the earnings from the selected stage, and a 2 EUR fixed payment for answering the questionnaire. On average, each subject earned 13 EUR, and each session lasted 90 minutes.

4 Experimental Results

4.1 Efficiency

One of the main arguments from proponents of affirmative action is that this policy does not lower the quality of either candidates or winners in competitions, and thereby does not harm efficiency. In our setup, we evaluate efficiency in terms of task performance (number of correct answers in the arithmetic task) of the candidate pool and of the selected candidates. We expect to confirm two findings from previous studies. First, affirmative action does not affect efficiency compared to no affirmative action in terms of the performance of either the

[8] Performance in the Arithmetic task and in the Grid task are moderately correlated (Pearson's correlation coefficient: $\rho = 0.412$).

[9] Chi Trieu, 2019. "Identity of affirmed groups and the consequences of affirmative action policies" December 02. Registry number AEARCTR-0005122.

candidate pool or the winners. Second, competition increases performance compared to no competition. We check the first point by comparing the performance of stage 3 winners and by comparing stage 1 and stage 3 performance of the candidate pools and the winners across treatments. To confirm the second point, we compare the average performance of all subjects among stages 1, 2, and 3.

Overall, we observe no significant differences in the task performance of either the candidate pools or the winners across treatments in both the compulsory competition and the self-selected competition. In all treatments, performance increases significantly when moving from the piece-rate incentive to the tournament incentive.[10],[11]

Result 1: Both gender quotas and mixed quotas do not cause efficiency loss compared to no quotas.

Result 1 reinforces a desirable feature of gender quotas that this policy does not affect efficiency measured by how winners are selected and their performance (Balafoutas and Sutter (2012); Balafoutas, Davis, and Sutter (2016); Niederle and Vesterlund (2007)). It further shows that this feature also holds true when quotas target two groups. It is noteworthy that in our setup, mixed quotas are implemented without the threat of sabotage or retaliation against affirmed individuals. With retaliation, Fallucchi and Quercia (2018) find that efficiency in terms of performance is hampered compared to without retaliation.

4.2 Selection into tournament

4.2.1 Willingness to compete at the aggregate level

We expect that subjects self-select into competition more when they are favored by affirmative action and vice-versa.

The left panel of Figure 1 shows the selection into competition for each gender in each treatment. We observe an encouraging effect of quotas on the willingness to compete of women in MIX, but not in GQ. The right panel of Figure 1 shows competition entry by types in CTR and MIX. The reaction of types in MIX closely corresponds to this incentive. Women with the color green have the highest and a significant rise in selection into tournament, followed by women with the color blue. There is a decrease in the competition

[10] See Appendix B, subsection B.1 for further details on the pairwise comparisons of performance across treatments. Throughout the paper, unless noted otherwise, we report the results of two-sided tests.

[11] We provide power calculations for the non-parametric tests for all of our main results in Appendix B, subsection B.7.

entry of men of both colors in MIX compared to the same types in CTR. Further, there is a discouraging effect of quotas on the competition entry of men.[12]

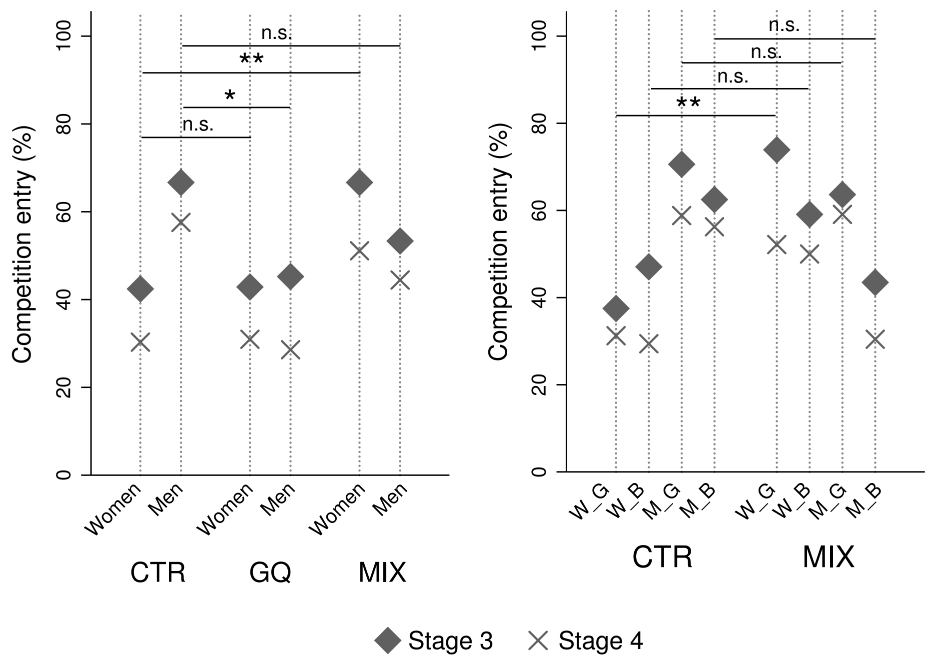

Figure 1: **Willingness to compete by treatment, gender and type**

Notes: The left panel displays the proportion of competition entry in stage 3 and stage 4 for each gender in each treatment. The right panel displays the proportion of competition entry in stage 3 and stage 4 for four types in treatment CTR and treatment MIX. The brackets and stars above each line show the results of Fisher's exact tests for competition entry in stage 3, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The absence of stars shows that a difference is not significant. Abbreviations: W_G - Women_Green, W_B - Women_Blue, M_G - Men_Green, M_B - Men_Blue.

Stage 4 accounts for effects of preference for performing. In this stage, subjects submit their stage 1 performance to either tournament or piece-rate incentive without having to perform. Asking subjects to submit their past performance to their preferred scheme accounts for preference for performing on top of preference for competing (Ifcher and Zarghamee (2016)). This can be measured by comparing competition entry of the same type in stage 3 and stage 4. Overall, competition entry in stage 4 is lower than in stage 3 in all treatments. The size of decrease is larger for women than for men in CTR and MIX, while it is larger for men than for women in GQ (see Figure 1). This evidence suggests that the presence or absence of performance does not necessarily result in variance in willingness to compete that is specific to gender.

[12]See Appendix B, subsubsection B.3.1 for details of the pairwise comparisons of women's and men's average willingness to compete across treatments.

4.2.2 Willingness to compete at the individual level

Table 1 shows the results of three probit regressions estimating the determinants of competition entry. All models estimate the effect of the treatments on the choice between tournament incentive and piece-rate incentive in stage 3. The predicted variable is a binary dummy, taking a value of 1 if a subject chooses the tournament incentive and 0 if a subject chooses the piece-rate incentive. We regress competition entry on treatment dummies (GQ and MIX) with CTR as the omitted group.

Model 1 regresses competition entry on treatment dummies with CTR as the base category and belief on own ranking in stage 3 as the control variable. In Model 2, we add risk attitude, two additional measures of overconfidence, and fairness perception of the policy in use as further controls. Model 3 estimates heterogeneous effects at the type level, adding an interaction term between the dummy for being favored and the treatment dummies.[13]

The first model shows that willingness to compete is largely driven by belief on own rank. On a scale from 1 to 6 with 1 as the best rank, one unit decrease in belief on rank lowers the likelihood of entering the tournament by 10.4 percentage points. Model 2 additionally shows that risk attitude is a further drive for contest entry. On an 11-point Likert scale, one unit increase in willingness to take risk increases the likelihood of competition entry by roughly 5 percentage points. Both gender quotas and mixed quotas have insignificant effects on competition entry in these models. When the indicator of being favored is included in Model 3, the result shows that mixed quotas significantly raise the probability of entering competition of the favored types by 34.0 percentage points. This effect under gender quotas is much lower, around 12.9 percentage points, and not significant.

Regarding control variables, we show that fairness perception and indirect measures of confidence matter for competition entry.[14] Results are robust if we use belief of rank in stage 2 instead of belief in stage 3, considering the argument that the decision to enter competition at the beginning of stage 3 is rather influenced by belief in the immediate previous stage 2.

Result 2a: A gender quota does not encourage women to enter competition while a mixed quota increases the competition entry of the targeted groups.

[13] Table B.1, Appendix B provides detailed results and descriptions of the control variables. In this table, we also report the results of a version of Model 3, with an interaction term between risk and the dummy for the favored types.

[14] For a more detailed analysis of fairness perception, see Appendix B, subsection B.2.

Table 1: Willingness to compete at the individual level

	(1)	(2)	(3)
GQ	-0.071	-0.007	-0.083
	(0.078)	(0.078)	(0.111)
MIX	0.043	0.100	-0.142
	(0.077)	(0.076)	(0.122)
GQ × Favored			0.129
			(0.146)
MIX × Favored			0.340**
			(0.148)
Favored			-0.084
			(0.110)
Belief of rank	-0.104***	-0.066***	-0.061***
	(0.020)	(0.022)	(0.022)
Risk measure		0.053***	0.054***
		(0.010)	(0.010)
Controls	No	Yes	Yes
N	240	240	240
Pseudo R^2	0.077	0.191	0.211

Notes: Average marginal effects of probit regressions estimating the likelihood of selecting the tournament incentive in stage 3. Standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The binary dependent variable is willingness to compete in stage 3 (1 if choosing tournament incentive, 0 if choosing piece-rate incentive). *Favored* is a dummy variable (1 if favored type, 0 otherwise). *Belief of rank* is a categorical variable about beliefs of own rank (between 1 and 6, with 1 as the best rank) in stage 3; *Risk measure* is the answer to the general risk question elicited on an 11-point scale, higher numbers indicate a higher willingness to take risks. Table B.1 provides detailed results and descriptions of the control variables.

4.2.3 Beliefs, self-confidence, and willingness to compete

What leads to the seemingly surprising ineffectiveness of gender quotas? Our data provide evidence that gender quotas lower belief of relative performance of both men and women.

As such, gender quotas fail to encourage women to compete more, at the cost of lowering the willingness to compete of men, in comparison with no gender quotas. More importantly, this low level of confidence is not backed up by an actual difference in both absolute and relative performance.[15]

In Table 2, we regress beliefs on treatment dummies. Model 1 controls for task performance in stage 3, and shows that gender quotas significantly increase belief (implying a worse perceived rank) by 0.324. When additional controls are taken into consideration in Model 2, the effect size increases to 0.405. Model 3 further disentangles this effect for the (non)favored types. Gender quotas significantly increase the beliefs of both men (by 0.700) and women (by 0.430). Mixed quotas increase the beliefs of the favored types by 0.104, and of the non-favored type by 0.143, yet both effects are not significant. Model 4 regresses belief on the interaction between treatment and gender, controlling for task performance in stage 3. Gender quotas significantly decrease the belief of men by 0.52 compared to CTR, and by 0.54 compared to MIX.[16] We observe a smaller size but significant effect when additional controls are accounted for in Model 5.[17]

These results suggest that gender quotas lead to lower self-confidence. Women and men under gender quotas are less confident in their relative performance, thereby becoming less willing to enter competition. In contrast, mixed quotas do not seem to be prone to such consequences for the targeted groups.

Result 2b: A gender quota leads to lower self-confidence of both men and women on their competence.

While the evidence above suggests a positive effect of mixed quotas on the willingness to compete of the targeted groups, it is important to interpret this result considering the potential signaling effect of the color quota in MIX. The implementation of affirmative actions favoring a certain group might carry a signal on the competence of this group. As such, gender quotas might lead to the belief that women are not as good as men at the arithmetic task. In MIX, in addition to a gender quota, there is a color quota based on a random split that likely carries no signal about the performance of the favored Green

[15]The non-parametric tests on the average belief on performance ranking and average performance of women across treatments do not always give significant results at the conventional levels. See Appendix B, subsection B.4 for further details.

[16]t-tests on the equality of the coefficients give p=0.030, p=0.042, and p=0.919 for the independent variables GQ x Men versus CTR x Men, GQ x Men versus MIX x Men, and CTR x Men versus MIX x Men respectively

[17]t-tests on the equality of the coefficients give p=0.013, p=0.024, and p=0.676 for the independent variables GQ x Men versus CTR x Men, GQ x Men versus MIX x Men, and CTR x Men versus MIX x Men respectively.

Table 2: The effect of quotas on beliefs

	(1)	(2)	(3)	(4)	(5)
GQ	0.324*	0.405***	0.700**	0.242	0.429
	(0.163)	(0.140)	(0.264)	(0.279)	(0.260)
MIX	-0.071	-0.045	0.143		
	(0.153)	(0.132)	(0.270)		
GQ × Favored (Women)			0.430*	0.130	0.160
			(0.240)	(0.297)	(0.254)
MIX × Favored			0.104		
			(0.192)		
Favored			0.270		
			(0.287)		
CTR × Men				-0.276	-0.270
				(0.313)	(0.313)
MIX × Women				-0.116	-0.160
				(0.238)	(0.190)
MIX × Men				-0.301	-0.154
				(0.292)	(0.269)
Stage 3 performance	-0.166***	-0.121***	-0.121***	-0.165 ***	-0.121***
	(0.018)	(0.020)	(0.021)	(0.019)	(0.021)
Constant	4.502***	4.649***	4.467***	4.63***	4.741***
	(0.221)	(0.286)	(0.354)	(0.279)	(0.296)
Controls	No	Yes	Yes	No	Yes
N	240	240	240	240	240
R^2	0.259	0.357	0.363	0.265	0.363

Notes: OLS estimation on effect of gender quotas and mixed quotas on belief on rank in stage 3. Clustered standard errors at the group level in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The dependent variable is a categorical variable about beliefs of own rank (between 1 and 6, with 1 as the best rank) in stage 3. *Favored* is a dummy variable (1 if favored type, 0 otherwise). Table B.2 provides detailed results and descriptions of the control variables.

group. The color quotas might thus mute the signaling effect that gender quotas might have if they were to be implemented alone. Subjects in MIX might infer that gender quotas, like color quotas, do not have any informational value about competency.

4.3 Post-competition cooperation

As subjects enter a new teamwork setup without affirmative action, post-competition team selection and team performance are likely to be affected by the exposure to affirmative action rules (or the lack thereof) in the previous stages.

We observe out-group favoritism regarding gender in CTR and in contrast, in-group favoritism regarding gender in MIX. A similar tendency holds for colleague selection. Women are selected significantly less than men in all treatments, with a larger gap in GQ and a smaller gap in MIX compared to CTR.[18] In terms of vote composition, out-group favoritism regarding gender exists in all treatments. Women receive more votes from men than from other women, while men receive more votes from women than from other men.

Result 3: Without affirmative action, male leaders are perceived as more competent than female leaders. Gender quotas worsen the perceived competence gap between male and female leaders, while mixed quotas lessen it.

Result 3 points to the unintentional "stigma of incompetence" outcome of gender quotas, where women under this policy are perceived as less competent by others (Heilman, Block, and Lucas (1992)). Given stage 5 setup, result 3 suggests that the difference in the presumed competence of women is based on priors formed by previous exposure to different treatments. In addition, the general tendency of favoring male leaders in all treatments might be because the arithmetic task is perceived as a typical "stereotypical-male" task. Comparing grid task performance of women and men across treatments to test whether such beliefs are precise or erroneous, we observe no significant difference.[19]

It is worth re-mentioning that the color quotas in MIX might have an informational effect on the perceived competency of the favored group that drives the team selection choices. Subjects in MIX might perceive gender quotas arbitrary as color quotas, thus choosing female leaders more often than subjects in GQ.

To study how team (mis)selection affects team performance. We calculate the deviation in percentage points between realized team performance and optimal team performance.[20] We observe no significant differences in overall inefficiency and in inefficiency due to team selection across treatments.[21]

[18] See Appendix B, subsection B.5 for details of the pairwise comparison with regards to the average number of votes for leaders and for colleagues by treatment. This section also presents the gender composition of leaders across treatments.

[19] See Appendix B, subsection B.6 for a detailed analysis.

[20] For example, efficiency loss due to leader selection is calculated as (1 - (Performance of chosen leader/Performance of the best leader)) × 100%

[21] See Appendix B, subsection B.6 for a detailed analysis.

In sum, gender quotas and mixed quotas do not harm cooperation measured as effort exerted in teamwork. However, efficiency loss emerges in all treatments if team misselection is considered, whereby mixed quotas generate the least inefficiency caused by leader misselection.

5 Conclusion

In an increasingly diverse population and workforce, institutions have adapted their affirmative action policies to target multiple disadvantaged groups. In many cases, more than one affirmative action policy is implemented. Although decision-makers are aware of how each policy influences competition outcomes, applicants are usually uninformed thereof. For applicants, this feature entails uncertainty about the actual favored group. We are the first to study how this feature affects outcomes of affirmative action policies in the form of quota rules in a laboratory experiment. We vary the rules of affirmative action, with one treatment favoring women with certainty and another treatment favoring either women or a member of one arbitrarily assigned group with a fifty-fifty probability. We compare both immediate and spillover outcomes of each rule with outcomes of a baseline treatment where affirmative action is not implemented.

We found that independent of the policy in use, efficiency measured by task performance is not affected. This result is in line with previous literature (Balafoutas and Sutter (2012); Niederle, Segal, and Vesterlund (2013); Balafoutas, Davis, and Sutter (2016)), confirming that affirmative action is not harmful to efficiency. Interestingly, gender quotas fail to encourage women to enter the competition. We argue that the mechanism at work is lower self-confidence. Gender quotas lower women's confidence in their relative performance, hence lowering their willingness to compete. In contrast, mixed quotas raise the tournament entry of both favored groups.

Regarding indirect outcomes, we find evidence that when affirmative action policy targets solely gender and is explicitly communicated as such, the policy seems to activate a stereotype threat about the competence of the favored group. Women are perceived as less competent in subsequent teamwork. The uncertainty about the favored group reverses this effect. Putting our findings in perspective with the related literature, negative consequences of gender quotas observed in previous studies (Leibbrandt, Wang, and Foo (2017); Ip, Leibbrandt, and Vecci (2020)) might capture an upper bound of unintended effect because women are the only favored group in these studies, and their social identity is made salient.

The implementation of group identities in our designs is useful for the analysis yet subject to certain limitations. In reality, there is no affirmative action implemented at random similar to the color quota in our design. Affirmative action in the real world is typically implemented based on criteria that identify the underrepresented groups, and as such carry a signaling effect on the performance of the targeted group. This signaling effect cannot be replicated by the color quota in our experiment. Under gender quota, subjects might perceive women as less good at performing the arithmetic task than men. However, this informational value might be muted when a color quotas are added to gender quotas, as subjects may now perceive both as arbitrary. To the extent that the "color" characteristic could resemble real-world affirmative actions, they would better resemble unobservable identities, such as socio-economic status or religion, but not other identities that are easy to make salient, such as age or disability. Last but not least, we study a rather specific case when both identities are favored by affirmative action. Further research with more complex group identities, e.g, additionally including a group without any favored characteristics could provide interesting insights.

Overall, our study serves as a thought experiment to study the likely effect of mixed quotas. Our results provide evidence that uncertainty over the targeted group and the salience of social identity do matter for both immediate and spillover outcomes of affirmative action policies. More generally, we illustrate how social identity is malleable under an institution, and in turn, affects the effectiveness of this institution. In organizational contexts, managers and policymakers might want to simultaneously target several disadvantaged groups and explicitly communicate affirmative action as such. At the same time, they might want to avoid strengthening the identity of selected candidates as being the favored or the unfavored.

References

Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman. 2011. "Reference Points and Effort Provision." *American Economic Review* 101 (2):470–92.

Ahern, Kenneth and Amy K. Dittmar. 2012. "The Changing of the Boards: The Impact on Firm Valuation of Mandated Female Board Representation." *The Quarterly Journal of Economics* 127 (1):137–197.

Almås, Ingvild, Alexander Cappelen, Kjell G Salvanes, Erik Sørensen, and Bertil Tungodden. 2016. "Willingness to Compete: Family Matters." *Management Science* 62 (8):2149–2162.

Araujo, Felipe A., Erin Carbone, Lynn Conell-Price, Marli W. Dunietz, Ania Jaroszewicz, Rachel Landsman, Diego Lamé, Lise Vesterlund, Stephanie W. Wang, and Alistair J. Wilson. 2016. "The slider task: an example of restricted inference on incentive effects." *Journal of the Economic Science Association* 2 (1):1–12.

Arcidiacono, Peter and Michael Lovenheim. 2016. "Affirmative Action and the Quality-Fit Trade-Off." *Journal of Economic Literature* 54 (1):3–51.

Bagde, Surendrakumar, Dennis Epple, and Lowell Taylor. 2016. "Does Affirmative Action Work? Caste, Gender, College Quality, and Academic Success in India." *American Economic Review* 106 (6):1495–1521.

Balafoutas, Loukas, Brent J. Davis, and Matthias Sutter. 2016. "Affirmative action or just discrimination? A study on the endogenous emergence of quotas." *Journal of Economic Behavior and Organization* 127:87–98.

Balafoutas, Loukas and Matthias Sutter. 2012. "Affirmative Action Policies Promote Women and Do Not Harm Efficiency in the Laboratory." *Science* 335 (6068):579–582.

———. 2019. "How uncertainty and ambiguity in tournaments affect gender differences in competitive behavior." *European Economic Review* 118 (C):1–13.

Banerjee, Ritwik, Nabanita Datta Gupta, and Marie Claire Villeval. 2018. "The spillover effects of affirmative action on competitiveness and unethical behavior." *European Economic Review* 101 (C):567–604.

———. 2020. "Feedback spillovers across tasks, self-confidence and competitiveness." *Games and Economic Behavior* 123:127 – 170.

Beaman, Lori, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova. 2009. "Powerful Women: Does Exposure Reduce Bias?" *The Quarterly Journal of Economics* 124 (4):1497–1540.

Beaurain, Guillaume and David Masclet. 2016. "Does affirmative action reduce gender discrimination and enhance efficiency? New experimental evidence." *European Economic Review* 90:350 – 362.

Bertrand, Marianne, Sandra E. Black, Sissel Jensen, and Adriana Lleras-Muney. 2018. "Breaking the Glass Ceiling? The Effect of Board Quotas on Female Labour Market Outcomes in Norway." *The Review of Economic Studies* 86 (1):191–239.

Calsamiglia, Caterina, Jörg Franke, and Pedro Rey-Biel. 2013. "The incentive effects of affirmative action in a real-effort tournament." *Journal of Public Economics* 98:15–31.

Chattopadhyay, Raghabendra and Esther Duflo. 2004. "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India." *Econometrica* 72 (5):1409–1443.

Chen, Zhuoqiong, David Ong, and Roman Sheremeta. 2015. "Competition Between and Within Universities: Theoretical and Experimental Investigation of Group Identity and the Desire to Win." *MPRA Working Paper* No. 67552.

Chowdhury, Subhasish M., Joo Young Jeon, and Abhijit Ramalingam. 2016. "Identity and group conflict." *European Economic Review* 90:107 – 121.

Corgnet, Brice, Roberto Hernán-González, and Eric Schniter. 2015. "Why real leisure really matters: incentive effects on real effort in the laboratory." *Experimental Economics* 18 (2):284–301.

Cornaglia, Francesca, Michalis Drouvelis, and Paolo Masella. 2019. "Competition and the role of group identity." *Journal of Economic Behavior & Organization* 162:136–145.

Council of European Union. 2000. "Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation." https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32000L0078.

Croson, Rachel and Uri Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic Literature* 47 (2):448–74.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde. 2018. "Global Evidence on Economic Preferences." *The Quarterly Journal of Economics* 133 (4):1645–1692.

Fallucchi, Francesco and Simone Quercia. 2018. "Affirmative action and retaliation in experimental contests." *Journal of Economic Behavior and Organization* 156:23 – 40.

Federal Law Gazette. 2015. "Gesetz für die gleichberechtigte Teilhabe von Frauen und Männern an Führungspositionen in der Privatwirtschaft und im öffentlichen Dienst (Law on Equal Participation of Women and Men in Leadership Positions in the Private and Public Sector)." Bundesgesetzblatt Jahrgang 2015 Teil I Nr.17 S. 642.

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

Frederick, Shane. 2005. "Cognitive Reflection and Decision Making." *Journal of Economic Perspectives* 19 (4):25–42.

Gächter, Simon, Lingbo Huang, and Martin Sefton. 2016. "Combining "real effort" with induced effort costs: the ball-catching task." *Experimental Economics* 19 (4):687–712.

Gee, Laura K. 2019. "The More You Know: Information Effects on Job Application Rates in a Large Field Experiment." *Management Science* 65 (5):2077–2094.

Gillen, Ben, Erik Snowberg, and Leeat Yariv. 2019. "Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study." *Journal of Political Economy* 127 (4):1826–1863.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini. 2003. "Performance in Competitive Environments: Gender Differences*." *The Quarterly Journal of Economics* 118 (3):1049–1074.

Goerg, Sebastian J., Sebastian Kube, and Jonas Radbruch. 2019. "The effectiveness of incentive schemes in the presence of implicit effort costs." *Management Science* 65 (9):3949–4450.

Greiner, Ben. 2004. "The Online Recruitment System ORSEE - A Guide for the Organization of Experiments in Economics." Papers on strategic interaction, Max Planck Institute of Economics, Strategic Interaction Group.

Gupta, Nabanita D., Poulsen Anders, and Marie Claire Villeval. 2013. "Gender Matching and Competitiveness: Experimental Evidence." *Economic Inquiry* 51 (1):816–835.

Heilman, Madeline E., Caryn J. Block, and Jonathan A. Lucas. 1992. "Presumed Incompetent? Stigmatization and Affirmative Action Efforts." *Journal of Applied Psychology* 77 (4):536–544.

Holzer, Harry and David Neumark. 2000. "Assessing Affirmative Action." *Journal of Economic Literature* 38 (3):483–568.

Hyman, Richard, Alain Klarsfeld, Eddy Ng, and Rana Haq. 2012. "Introduction: Social regulation of diversity and equality." *European Journal of Industrial Relations* 18 (4):279–292.

Ibanez, Marcela and Gerhard Riener. 2018. "Sorting through Affirmative Action: Three Field Experiments in Colombia." *Journal of Labor Economics* 36 (2):437–478.

Ifcher, John and Homa Zarghamee. 2016. "Do gender-variant preferences for competition persist in the absence of performance?" *Economic Inquiry* 54 (4):1918–1930.

Ip, Edwin, Andreas Leibbrandt, and Joseph Vecci. 2020. "How Do Gender Quotas Affect Workplace Relationships? Complementary Evidence from a Representative Survey and Labor Market Experiments." *Management Science* 66 (2):503–1004.

Iriberri, Nagore and Pedro Rey-Biel. 2019. "Competitive Pressure Widens the Gender Gap in Performance: Evidence from a Two-stage Competition in Mathematics." *Economic Journal* 129 (620):1863–1893.

Jensenius, Francesca Refsum. 2015. "Development from Representation? A Study of Quotas for the Scheduled Castes in India." *American Economic Journal: Applied Economics* 7 (3):196–220.

Kato, Takao and Pian Shu. 2016. "Competition and social identity in the workplace: Evidence from a Chinese textile firm." *Journal of Economic Behavior & Organization* 131 (Part A):37–50.

Kölle, Felix. 2017. "Affirmative action, cooperation, and the willingness to work in teams." *Journal of Economic Psychology* 62:50 – 62.

Lane, Tom. 2016. "Discrimination in the laboratory: A meta-analysis of economics experiments." *European Economic Review* 90:375 – 402. Social identity and discrimination.

Leibbrandt, Andreas and John A List. 2018. "Do Equal Employment Opportunity Statements Backfire? Evidence From A Natural Field Experiment On Job-Entry Decisions." *NBER Working Paper* (25035).

Leibbrandt, Andreas, Liang Choon Wang, and Cordelia Foo. 2017. "Gender quotas, competitions, and peer review: Experimental evidence on the backlash against women." *Management Science* 64 (8):3501–3516.

Leonard, Kenneth, John List, and Uri Gneezy. 2009. "Gender Differences in Competition: Evidence From a Matrilineal and a Patriarchal Society." *Econometrica* 77:1637–1664.

Maggian, Valeria and Natalia Montinari. 2017. "The spillover effects of gender quotas on dishonesty." *Economics Letters* 159 (C):33–36.

Matsa, David A. and Amalia R. Miller. 2013. "A Female Style in Corporate Leadership? Evidence from Quotas." *American Economic Journal: Applied Economics* 5 (3):136–69.

Mollerstrom, Johanna. 2022. "Favoritism and cooperation." *Public Choice* 191:293–307.

Niederle, Muriel, Carmit Segal, and Lise Vesterlund. 2013. "How costly is diversity? affirmative action in light of gender differences in competitiveness." *Management Science* 59 (1):1–16.

Niederle, Muriel and Lise Vesterlund. 2007. "Do Women Shy Away From Competition? Do Men Compete Too Much?" *The Quarterly Journal of Economics* 122 (3):1067–1101.

———. 2011. "Gender and Competition." *Annual Review of Economics* 3 (1):601–630.

Pande, Rohini. 2003. "Can Mandated Political Representation Increase Policy Influence for Disadvantaged Minorities? Theory and Evidence from India." *American Economic Review* 93 (4):1132–1151.

Petters, Lea and Marina Schroeder. 2020. "Negative Side Effects of Affirmative Action: How Quotas Lead to Distortions in Performance Evaluation." *European Economic Review* 130:103500.

Schildberg-Hörisch, Hannah, Marco Schwarz, Chi Trieu, and Jana Willrodt. 2023. "Perceived fairness and consequences of affirmative action policies." *forthcoming in Economic Journal* .

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

SGB IX – *German Social Code, Volume 9*. 2001. "§ 154 SGB IX Pflicht der Arbeitgeber zur Beschäftigung schwerbehinderter Menschen (Employers' duty of employing people with disabilities)." Https://www.sozialgesetzbuch-sgb.de/sgbix/154.html. Accessed May 22, 2020.

Sowell, Thomas. 2004. *Affirmative Action around the World: An Empirical Study.* Yale UniversityPress.

Sutter, Matthias and Daniela Glätzle-Rützler. 2015. "Gender Differences in the Willingness to Compete Emerge Early in Life and Persist." *Management Science* 61 (10):2339–23354.

Sutter, Matthias, Claudia Zoller, and Daniela Glätzle-Rützler. 2019. "Economic behavior of children and adolescents – A first survey of experimental economics results." *European Economic Review* 111 (C):98–121.

Wagner, Gert G, Joachim R Frick, and Jürgen Schupp. 2007. "The German Socio-Economic Panel Study (SOEP) - evolution, scope and enhancements." *SOEPpapers on Multidisciplinary Panel Data Research, No. 1* .

A Additional related literature

As mentioned in section 1 (Introduction) of the main paper, favoring one particular minority seems subjective and might fail to fully capture the truly disadvantaged individuals in an increasingly diverse population and workforce. Take gender quotas for example, several studies reveal the potential inefficiency when classical separation of advantaged and disadvantaged groups are applied. For example, men also shy away from tournaments in team competition (Dargnies (2012)). Apart from gender, other factors also influence willingness to compete, such as socio-economic backgrounds (Almas et al. (2016)), age, context, and type of task (Dreber, von Essen, and Ranehill (2011, 2014)).

As also stated in section 1 (Introduction), our experiment contributes in two ways: (i) we introduce uncertainty about the favored group, thus varying the salience of the social identity of the beneficiaries across treatments, and (ii) we add perceived competence as a new measure of the spillover outcomes of affirmative action. Regarding the second contribution, our motivation to measure perceived competence is twofold. First, a fair amount of studies show that performance in real-effort tasks with a short working time implemented in laboratory experiments is difficult to influence. In such settings, subjects usually exert maximal effort.[22] This tendency might mask the null result that affirmative action does not harm cooperation observed in previous studies (Balafoutas, Davis, and Sutter (2016); Kölle (2017)) in which cooperation is measured by effort provision. Our design introduces additional measures of cooperation that are not subject to such constraints. Second, team selection most likely occurs in real-life workplaces, where many projects allow employees to self-select into teams with colleagues. In this scenario, a comprehensive measure of spillover effects of affirmative action is not only effort provision but also the structure of teams and the efficiency gain (loss) in terms of team performance, if the team members had been selected differently.

Gender quotas have been examined in numerous settings such as laboratory experiments (Balafoutas and Sutter (2012); Niederle, Segal, and Vesterlund (2013); Balafoutas, Davis, and Sutter (2016); Beaurain and Masclet (2016); Kölle (2017); Leibbrandt, Wang, and Foo (2017); Maggian and Montinari (2017); Ip, Leibbrandt, and Vecci (2020)), field experiments (Chattopadhyay and Duflo (2004); Ibanez and Riener (2018); Beaman et al. (2009)), and empirical studies on mandatory quotas on corporate boards in Norway (Ahern and Dittmar (2012); Matsa and Miller (2013); Bertrand et al. (2018)). Quotas targeting castes in India (Pande (2003); Jensenius (2015); Bagde, Epple, and Taylor (2016); Banerjee, Gupta, and

[22] See, e.g, Corgnet, Hernán-González, and Schniter (2015), Araujo et al. (2016), Gächter, Huang, and Sefton (2016), Goerg, Kube, and Radbruch (2019).

Villeval (2018)) and students of ethnic minorities in the U.S. (see, e.g, Arcidiacono and Lovenheim (2016)) have also been extensively researched.

Apart from quotas, other forms of studied affirmative action include bonuses and lump sum payments (Schotter and Weigelt (1992); Calsamiglia, Franke, and Rey-Biel (2013); Balafoutas and Sutter (2012)), repetition of competition conditional on a threshold of positions being filled by the favored group (Balafoutas and Sutter (2012)), reservation of an extra prize for the targeted group (Fallucchi and Quercia (2018)), and statements of equal employment opportunities (Leibbrandt and List (2018)).

Outcomes of affirmative action policies documented in experimental studies are by far ambiguous. On one hand, gender quotas increase competition entry of women, at the same time do not distort men's willingness to compete and do not harm efficiency (Balafoutas and Sutter (2012), Niederle, Segal, and Vesterlund (2013), Balafoutas, Davis, and Sutter (2016). Beaurain and Masclet (2016) find that gender quotas improve subjective evaluation of women's competence in a hiring game, thus increasing their employment. Quotas for castes enhance confidence and thereby raise willingness to compete of the favored group (Banerjee, Gupta, and Villeval (2020)). Regarding spillover effects, gender quotas do not result in subsequent unethical behaviors (Maggian and Montinari (2017); Banerjee, Gupta, and Villeval (2018)) or spiteful behavior against the affirmed group (Banerjee, Gupta, and Villeval (2018)). Neither is this policy harmful to effort provision in teamwork (Balafoutas, Davis, and Sutter (2016), Kölle (2017)), willingness to work in team (Kölle (2017)), or coordination with team members (Balafoutas and Sutter (2012)).

On the other hand, gender quotas might backfire and generate sabotage against women, consequently discouraging them from entering competition (Leibbrandt, Wang, and Foo (2017)). Affirmative action in the form of an equal employment opportunity statement is found to discourage rather than empower the racial minority in job applications (Leibbrandt and List (2018)). When the favored group is randomly assigned, retaliation threat is detrimental to the competition entry of the favored group (Fallucchi and Quercia (2018)), and a quota reduces cooperation in following public good games (Mollerstrom (2022)).

Our study is related to several real-world policies. Many institutions have reformulated their diversity management policies to target multiple marginalized groups. For example, the European Union recognized the importance of anti-discrimination measures not only for gender and ethnic minorities, but also regarding age, disability, religion, and sexual orientation (Hyman et al. (2012)), thus mandating member states to address discrimination in these areas in employment and occupation (Council of European Union (2000)). In Germany, since 2001 employers with more than 20 employees have had to reserve at least 5% of positions for individuals with disabilities (SGB IX – *German Social Code, Volume 9*

(2001)) and since 2016, at least 30% of the supervisory or administrative boards of publicly listed companies have to be filled by the underrepresented gender (Federal Law Gazette (2015)). In India, seats at higher education institutions are reserved through two quota systems, one for women and one for members of certain disadvantaged social groups (castes) (Bagde, Epple, and Taylor (2016)).

One contribution of this paper, namely investigating outcomes of affirmative action policies when there are two targeted groups, and the affirmed group is uncertain differs from previous studies in the same vein. Bagde, Epple, and Taylor (2016) and Leibbrandt and List (2018) study affirmative action policies for more than one group of beneficiaries. Bagde, Epple, and Taylor (2016) study quotas for castes and women in university admission using data from India. Outcomes of quotas are evaluated through college attendance and the quality of selected colleges of the targeted groups. Leibbrandt and List (2018) investigate the impact of a statement emphasizing equal opportunity regarding gender, racial minorities, age, any other protected characteristics on job entry. Compared to these studies, our research differs significantly in the outcomes investigated, and uncertainty and salience of social identity as the main aspect of interest.

B Additional Results

B.1 Efficiency

In the compulsory competition, we observe no significant differences in the task performance of winners across treatments. In CTR, winners solve an average of 13.68 tasks correctly, while this number in GQ is 14.32 and in MIX it is 14.5 (two-sided Mann-Whitney U test, CTR vs. GQ: p= 0.568, CTR vs. MIX: p= 0.356, and GQ vs. MIX: p= 0.638).[23] A Kruskal-Wallis test does not detect any significant differences across all treatments (p = 0.631).[24]

In the self-selected competition, Figure B.1 shows the task performance in stage 1 and stage 3 of subjects who choose tournament incentive in stage 3 in the left panel, and of winners in the right panel. Again, we do not observe any significant differences. The null hypotheses on the equality of means cannot be rejected for any pairwise comparisons of treatments (test results reported in the figure note).

[23]Throughout the paper, unless noted otherwise, we report the results of two-sided tests.

[24]We provide power calculations for the non-parametric tests regarding our main results in subsection B.7.

Extending the analysis to performance of all subjects in stage 1, we obtain a similar result that performance does not significantly differ across treatments, confirming the validity of our sampling (see also Figure B.2).

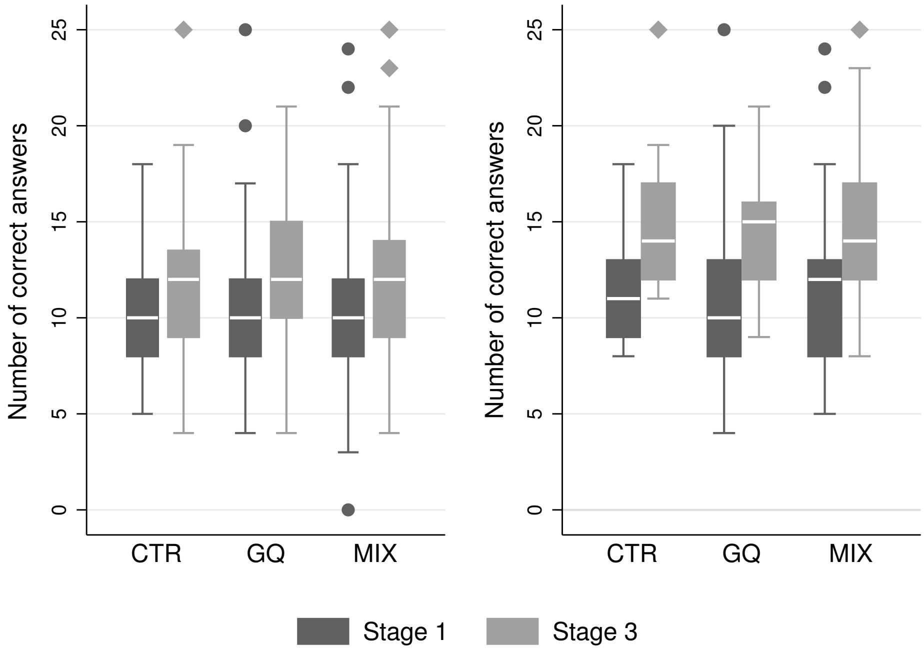

Figure B.1: **Performance in stage 1 and stage 3 of subjects self-selecting into competition (left) and of winners in stage 3 (right).**

Notes: The figure displays performance under piece-rate and tournament incentives of subjects who self-select into tournament (left panel) and of winners (right panel) in stage 3. The upper (lower) hinges of the boxes show the 75th (25th) percentiles, the white lines inside the boxes show the median values, the upper (lower) adjacent lines show the maximum (minimum), and the points show outliers. No significant difference in efficiency across treatments. The average performance of subjects opting for tournament in stage 1 (stage 3) in CTR is 10.22 (11.69), in GQ is 10.54 (12.16) and in MIX is 10.24 (12.19). Mann-Whitney U test, CTR vs. GQ: p= 0.903 (p= 0.568), CTR vs. MIX: p= 0.885 (p= 0.709), GQ vs. MIX: p= 0.782 (p= 0.827) for performance in stage 1 (stage 3). Kruskal-Wallis tests, p= 0.962 (p=0.852) for performance in stage 1 (stage 3). The average performance of stage 3 winners in stage 1 (stage 3) in CTR is 11.44 (14.88), in GQ is 11.30 (14.35) and in MIX is 11.77 (14.46). Mann-Whitney U test, CTR vs. GQ: p= 0.597 (0.761), CTR vs. MIX: p= 0.907 (p= 0.715), GQ vs. MIX: p= 0.547 (p= 0.894) for performance in stage 1 (stage 3). Kruskal-Wallis tests, p= 0.802 (p=0.925) for performance in stage 1 (stage 3).

In addition, our data provide evidence of the efficiency-enhancing effect of competition (Balafoutas, Davis, and Sutter (2016); Niederle and Vesterlund (2007) and Figure B.2). In all treatments, performance increases significantly when moving from piece-rate incentive to tournament incentive (Wilcoxon signed-rank test, $p < 0.001$ for all treatments). Under both incentives, the average performance of women is not significantly lower than men. In stage 1, women solve an average of 9.04 tasks correctly, while men score 10.08. In stage 2, the performance of women increases to 10.02 and the performance of men rises to 11.19 (Mann-Whitney U test, p= 0.202, p= 0.125 for performance in stage 1 and performance in stage 2 respectively).

Figure B.2 displays performance under piece-rate incentives and tournament incentives by gender and treatment. In stage 1, women (men) solve on average 8.73 (9.91) tasks correctly in CTR compared to 8.93 (10.17) GQ and 9.37 (10.13) in MIX. In stage 2, women (men) solve on average 10.03 (11.06) tasks correctly in CTR compared to 9.88 (11.55) GQ and 10.62 (10.96) in MIX. The corresponding p-values of Mann-Whitney U tests comparing performance between men and women are p= 0.128, p= 0.253, p= 0.939 for performance in stage 1, p= 0.238, p= 0.105, p= 0.997 for performance in stage 2 for CTR, GQ and MIX, respectively.

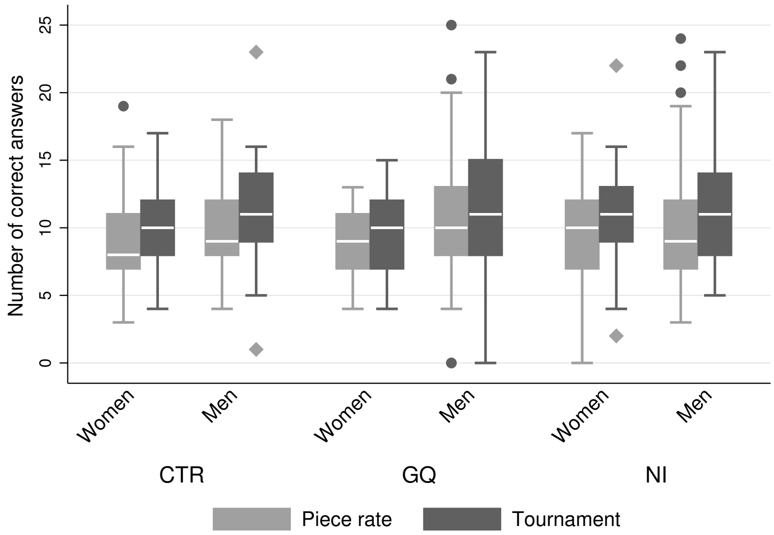

Figure B.2: **Performance in stage 1 and stage 2 by treatment and gender**

Notes: Boxplots of performance in stage 1 (piece-rate) and state 2 (tournament) by gender and treatment. The upper (lower) hinges of the boxes show the 75th (25th) percentiles, the white lines inside the boxes show the median values, the upper (lower) adjacent lines show the maximum (minimum) and the dots show outliers.

B.2 Fairness perception

The left panel of Figure B.3 shows the average perceived fairness for three policies under study by gender and treatment. In all treatments, women generally perceive gender quotas as fairer than men do (Mann-Whitney U test comparing average perceived fairness for gender quotas between men and women, p= 0.083, $p < 0.001$, $p < 0.001$ for CTR, GQ and MIX respectively). The gender gap in fairness perception for gender quotas is the largest in GQ. On one hand, women in GQ perceive gender quotas as fairer than women in other treatments, suggesting a self-serving bias (Mann-Whitney U test comparing the average perceived fairness for gender quotas among women across treatments, GQ vs. CTR:

p= 0.051, GQ vs. MIX: p= 0.469, and CTR vs. MIX: p= 0.187). On the other hand, men in GQ perceive this policy as less fair compared to men in other treatments (Mann-Whitney U test comparing the average perceived fairness for gender quotas among men across treatments, GQ vs. CTR: p= 0.895, GQ vs. MIX: p= 0.847, and CTR vs. MIX: p= 0.905). Intuitively, women as the targeted group of gender quotas in treatment GQ might anticipate that this policy is viewed unfavorably by men and consider their advantage as "unjustified", thus being reluctant to enter competition.

The right panel of Figure B.3 shows the average perceived fairness for three policies under study for the (non)favored types in MIX and their reference types (same types) in CTR and GQ. The gap in fairness perception for mixed quotas between the favored and the nonfavored types is not the largest in MIX, but it is between the reference types in GQ, suggesting that mixed quotas are relatively "accepted" by the nonfavored type in MIX.

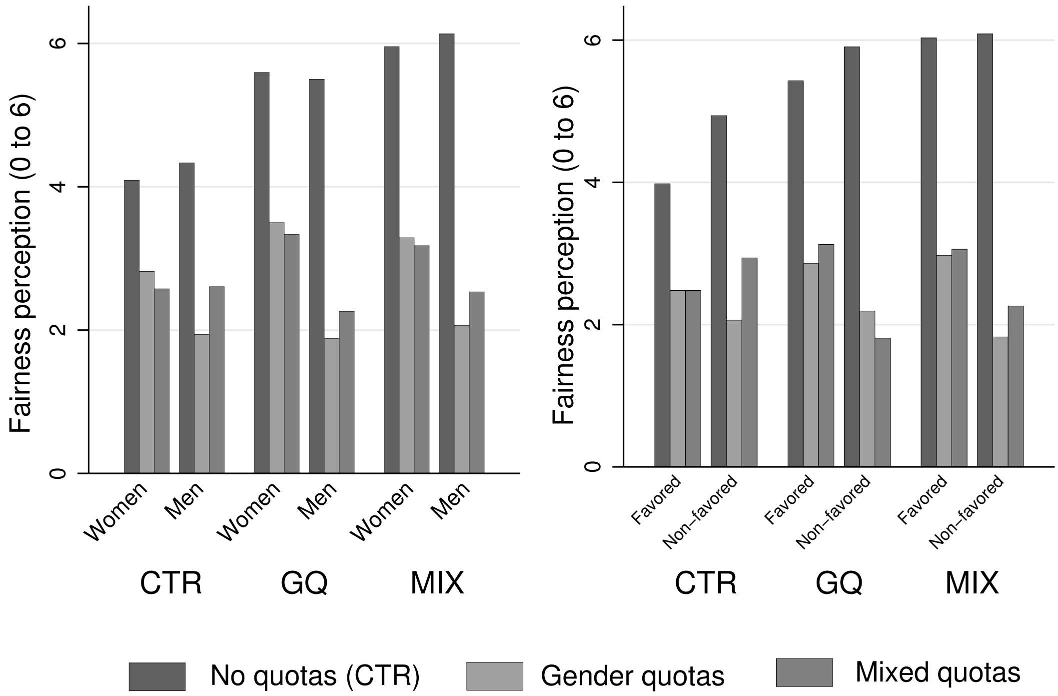

Figure B.3: **Fairness perception**

Notes: The left panel of the figure displays the average fairness perceptions for different policies by gender and treatment. The right panel of the figure displays the average perceived fairness for three policies under study for the (non)favored types in MIX and their reference types in CTR and GQ. The favored types are Women_Blue, Women_Green, Men_Green; and the nonfavored type is Men_Blue. Higher numbers indicate that a policy is perceived as fairer.

B.3 Willingness to compete

B.3.1 Willingness to compete at the aggregate level

In this subsection, we present details of the pairwise comparisons of women's and men's average willingness to compete across treatments to substantiate subsection 4.2 of the main paper.

First, we observe an encouraging effect of quotas on the willingness to compete of women in MIX, but not in GQ. The proportion of women selecting the tournament incentive is highest in MIX, and almost the same in GQ and in CTR. Under gender quotas, 42.9% of women choose tournaments, only 0.5 percentage points higher than the competition entry rate of women in CTR (Fisher's exact test, p= 1.000). Under mixed quotas, 66.7% of women select into the tournament, 24.3 percentage points significantly higher than in CTR (Fisher's exact test, p= 0.040) and as high as the willingness to compete of men in CTR. Men opt for competition more than women per se. In CTR, 66.7% (57.6%) of men choose tournament in stage 3 (stage 4), significantly more than the proportion of women at 42.4% (30.3%)(Fisher's exact test, p= 0.041 (0.046)). This evidence is in line with the findings of previous studies about the gender gap in the willingness to compete (Gneezy, Niederle, and Rustichini (2003), Niederle and Vesterlund (2007), Croson and Gneezy (2009), Leonard, List, and Gneezy (2009), Niederle and Vesterlund (2011), Sutter and Glätzle-Rützler (2015), Almås et al. (2016), Iriberri and Rey-Biel (2019)).

Second, there is a discouraging effect of quotas on the competition entry of men. In GQ, 45.2% of men choose tournaments, 21.5 percentage points lower than the competition entry rate of men in CTR (Fisher's exact test, p= 0.101). Men in MIX are also discouraged from selecting competitive incentives compared to CTR, but less so compared to GQ. Their competition entry rate is 53.3%, 13.4 percentage points lower than in CTR (Fisher's exact test, p= 0.255) and 8.1 percentage points higher than men in GQ (Fisher's exact test, p= 0.522).

B.3.2 Willingness to compete at the individual level

Table B.1: Willingness to compete at the individual level - details

	(1)	(2)	(3)	(4)
GQ	-0.072	-0.007	-0.083	-0.085
	(0.079)	(0.078)	(0.111)	(0.112)
MIX	0.043	0.100	-0.142	-0.142
	(0.077)	(0.076)	(0.122)	(0.122)
GQ × Favored			0.129	0.130
			(0.146)	(0.146)
MIX × Favored			0.340**	0.342**
			(0.148)	(0.147)
Favored			-0.084	0.036
			(0.110)	(0.188)
Belief of rank	-0.104***	-0.066***	-0.061***	-0.063***
	(0.020)	(0.022)	(0.022)	(0.022)
Risk measure		0.053***	0.054***	0.066***
		(0.010)	(0.010)	(0.018)
Risk measure × Favored				-0.020
				(0.025)
Fairness perception		0.046***	0.037**	0.038**
		(0.017)	(0.018)	(0.018)
Belief of CRT score		0.098**	0.109**	0.104**
		(0.046)	(0.046)	(0.046)
Belief on math ability		0.021*	0.022**	0.022**
		(0.011)	(0.011)	(0.011)
N	240	240	240	240
Pseudo R^2	0.077	0.191	0.211	0.213

Notes: Average marginal effects of probit regressions estimating the likelihood of selecting tournament incentive in stage 3. Standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The binary dependent variable is willingness to compete in stage 3 (1 if tournament, 0 if piece-rate). *Favored* is a dummy variable (1 if favored type, 0 otherwise). *Belief of rank* is a categorical variable about beliefs of own rank (between 1 and 6, with 1 as the best rank) in stage 3; *Risk attitude* is the answer to the general risk question elicited on an 11-point scale, higher numbers indicate a higher willingness to take risks; *Fairness perception* reflects fairness rating of own treatment, elicited on a 7-point scale on which higher numbers indicate higher perceived fairness; *Belief of CTR score* is belief of own rank in the Cognitive Reflection Test (Frederick (2005)), higher value indicates higher rank; *Belief on math ability* is the agreement on the statement "I am good at math", elicited on a 11-point scale, higher value indicates higher degree of agreement.

In relation to the control variables for risk attitude and confidence, Gillen, Snowberg, and Yariv (2019) replicates the experiment of Niederle and Vesterlund (2007) and show that the gender gap in willingness to compete from this design is prone to measurement error. When several measures of risk attitude and confidence are added to the estimation, this gap becomes insignificant. We take their findings into consideration and elicit two measures of risk preference and three measures of confidence. Two measures of risk preference include a qualitative measure and an incentivized choice list. Three measures of confidence include belief of rank in the arithmetic task, belief of performance in the Cognitive Reflection Test (Frederick (2005)), and perception of mathematical ability. We refer readers to Appendix C and Appendix E for the detailed measures.

B.4 The effects of quotas on beliefs

On a scale from 1 to 6, with 1 as the best rank, the average belief in stage 3 of women in GQ is 3.10, higher than the average belief of women in CTR at 2.97 (Mann-Whitney U test, p= 0.504), significantly higher than the average belief of the favored types in MIX at 2.61 (Mann-Whitney U test, p= 0.053), and marginally significantly higher than the average belief of women in MIX at 2.67 (Mann-Whitney U test, p= 0.105). The average belief in stage 3 of men in GQ is 2.93 in GQ, compared with 2.42 of men in CTR (Mann-Whitney U test, p= 0.189) and 2.53 of men in MIX (Mann-Whitney U test, p= 0.279).[25]

In terms of performance, women in GQ solve on average 10.12 tasks correctly while in MIX favored subjects solve an average of 11.12 tasks correctly in stage 3 (Mann-Whitney U test, p= 0.358), and women solve an average of 11.22 tasks correctly (Mann-Whitney U test, p= 0.174).

As women perform slightly better in MIX relative to GQ, it might be expected that they also rank themselves slightly better in MIX than in GQ. However, Table B.2 below (which provides further details for the regression results presented in Table 2 of the main paper) provides supporting evidence that the difference in confidence does not correctly correspond to the differences in performance.

[25] One might argue that subjects base their decision to enter the competition at the beginning of stage 3 on their belief on rank in the immediate previous stage 2. Replicating this analysis for belief on rank in stage 2, we observe a quite similar result that women in GQ is significantly less confident than favored subjects in MIX (Mann-Whitney U test, p= 0.070).

Table B.2: The effect of quotas on beliefs

	(1)	(2)	(3)	(4)	(5)
GQ	0.324*	0.405***	0.700**	0.242	0.429
	(0.163)	(0.140)	(0.264)	(0.279)	(0.260)
MIX	-0.071	-0.045	0.143		
	(0.153)	(0.132)	(0.270)		
GQ × Favored (Women)			0.430*	0.130	0.160
			(0.240)	(0.297)	(0.254)
MIX × Favored			0.104		
			(0.192)		
Favored			0.270		
			(0.287)		
CTR × Men				-0.276	-0.270
				(0.313)	(0.313)
MIX × Women				-0.116	-0.160
				(0.238)	(0.190)
MIX × Men				-0.301	-0.154
				(0.292)	(0.269)
Stage 3 performance	-0.166***	-0.121***	-0.121***	-0.165 ***	-0.121***
	(0.018)	(0.020)	(0.021)	(0.019)	(0.021)
Fairness perception		0.046	0.064		0.063
		(0.039)	(0.048)		(0.051)
Like task		-0.172***	-0.178***		-0.178***
		(0.044)	(0.042)		(0.044)
Belief on math ability		-0.064*	-0.062*		-0.062*
		(0.034)	(0.034)		(0.034)
Field of study		0.115**	0.112**		0.112**
		(0.052)	(0.053)		(0.053)
Constant	4.502***	4.649***	4.467***	4.63***	4.741***
	(0.221)	(0.286)	(0.354)	(0.279)	(0.296)
N	240	240	240	240	240
R^2	0.259	0.357	0.363	0.265	0.363

Notes: OLS estimation on effect of affirmative action on belief of rank in stage 3. Clustered standard errors at the group level in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The dependent variable is a categorical variable about beliefs of own rank (between 1 and 6, with 1 as the best rank) in stage 3. *Favored* is a dummy variable (1 if favored type, 0 otherwise); *Fairness perception* reflects fairness rating of own treatment, elicited on a 7-point scale, higher numbers indicate higher fairness; *Like task* is the answer to the question "How much do you like the arithmetic task?" elicited on an 7-point scale, higher numbers indicate higher liking; *Belief in math ability* is the agreement on the statement "I am good at math", elicited on a 11-point scale, higher value indicates higher degree of agreement.

B.5 Team selection

Tendency of favoring male leaders over female leaders in all treatments: Figure B.4 shows the average number of votes for leaders (left panel) and colleagues (right panel) by gender and treatment. We observe an overall tendency to favor male leaders over female leaders in all treatments.

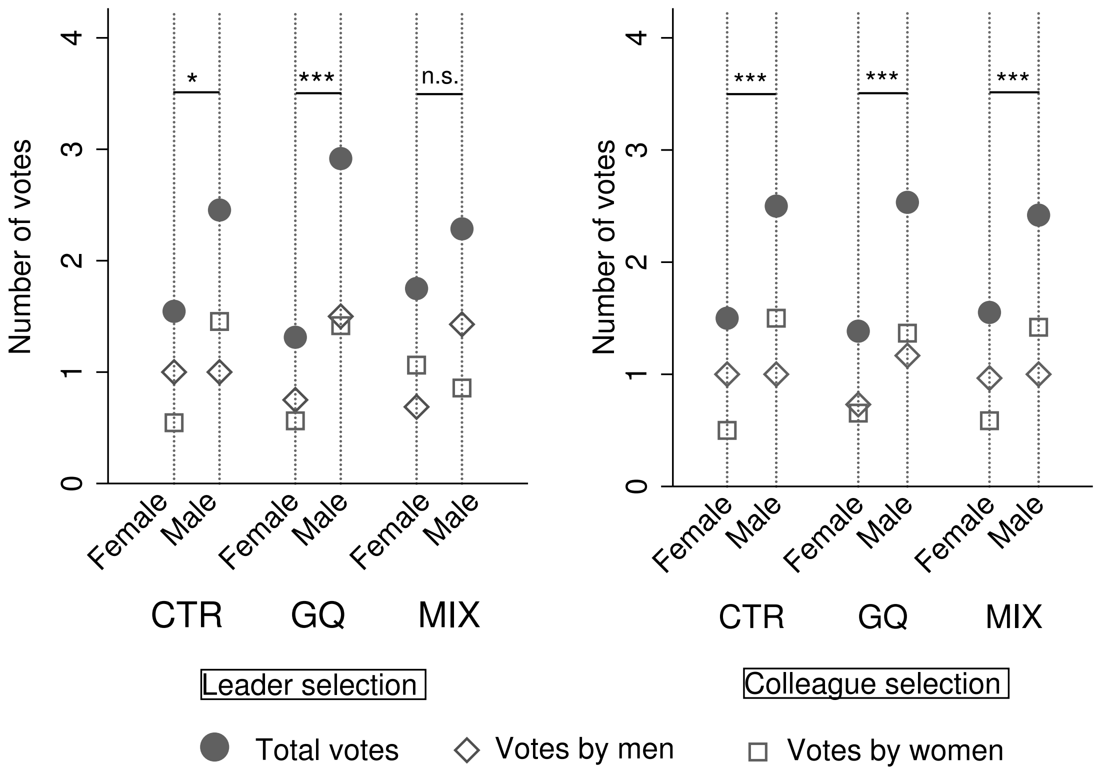

Figure B.4: **Team selection**

Notes: the figure displays the average number of votes for leaders (left panel) and non-leaders (right panel) in stage 5. Female (Male) leaders receive 1.55 (2.45), 1.31 (2.92), 1.75 (2.29) votes in CTR, GQ and MIX, respectively (Mann-Whitney U test comparing votes for female versus male leaders, p= 0.099, p= 0.002, p= 0.142 for CTR, GQ and MIX respectively). Female (Male) colleagues receive 1.50 (2.50), 1.38 (2.53) 1.55 (2.42) votes in CTR, GQ and MIX, respectively (Mann-Whitney U test comparing votes for female versus male non-leaders, $p < 0.001$ for all treatments). The brackets and stars above each line show results of Mann-Whitney U test, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The absence of stars shows that a difference is not significant.

In CTR, female leaders receive on average 1.55 votes while male leaders receive an average of 2.45 votes (Mann-Whitney U test, p= 0.099). This gap is *bigger and significant* in GQ while *smaller and insignificant* in MIX. On average, there are 1.31 votes for female leaders and 2.92 votes for male leaders in GQ (Mann-Whitney U test, p= 0.002); 1.75 votes for female leaders and 2.29 votes for male leaders in MIX (Mann-Whitney U test, p= 0.142). The difference in votes for male (female) leaders across treatments does not reach statistical significance.

In MIX, leaders from the Green (favored) group receive 0.6 votes on average, compared to 1.80 votes for leaders from the Blue (non-favored) group (Mann-Whitney U test, p= 0.272), and 2.25 (1.75) votes for the Green group in CTR (GQ) (Mann-Whitney U tests, CTR vs. MIX: p= 0.879, GQ vs. MIX: p=0.238).

Composition of votes: we observe out-group favoritism regarding gender in CTR. Female leaders receive on average 1.8 times more votes from male than female group members, while this pattern reverses for male leaders who receive on average 1.5 times more votes from female compared to male group members. The change in the average number of votes for both male and female leaders between CTR and GQ is largely driven by changes in votes by male group members. In particular, female (male) leaders in CTR receive 1.3 (1.5) times more (less) votes from male group members compared to female (male) leaders in GQ. Cross-gender voting decreases but does not completely disappear in GQ. In contrast, we observe in-group favoritism regarding gender in MIX. Female leaders in MIX receive more votes from women than men. On average, they receive 1.9 times more votes from female group members compared to female leaders in CTR. Male leaders in MIX receive 1.4 times more votes from male and 1.7 times fewer votes from female group members compared to male leaders in CTR.

Composition of leaders: To obtain a clearer picture, we take the gender composition of leaders into consideration. Table B.3 describes this composition for each treatment. We focus on groups with one leader of each gender, which account for 63.64%, 85.71%, and 80.00% of the sample size in CTR, GQ, and MIX respectively. This is the most pointed case as each non-leader chooses between a male leader and a female leader. For these groups, the pattern observed in the whole sample becomes even more prominent in terms of size. Female leaders are voted for less frequently than male leaders in CTR. This gap is wider in GQ and narrower in MIX (see Figure B.5 for specific pairwise comparisons).

Table B.3: Number of observations and gender composition of leaders by treatment.

	CTR		Gender quotas		Mixed quotas		Total
Two male leaders	12	(18.18%)	0	(0%)	6	(6.67%)	18
Two female leaders	12	(18.18%)	12	(14.29%)	12	(13.33%)	36
One female and one male leader	42	(63.64%)	72	(85.71%)	72	(80.00%)	186
Total	66		84		90		240

Figure B.5 displays the average number of votes for leaders (left panel) and non-leaders (right panel) in stage 5 for groups with one female leader and one male leader. Female (Male) leaders receive 1.29 (2.71), 1.08 (2.92), 1.67 (2.33) votes in CTR, GQ, and MIX, respectively (Mann-Whitney U test comparing votes for female versus male leaders, p= 0.023, p= 0.002, p= 0.050 for CTR, GQ, and MIX respectively). Female (Male) non-leaders receive 1.50 (2.50), 1.38 (2.63) 1.50 (2.50) votes in CTR, GQ, and MIX, respectively (Mann-Whitney U test comparing votes for female versus male non-leaders, p= 0.060, p= 0.002, p= 0.006 for CTR, GQ, and MIX respectively).

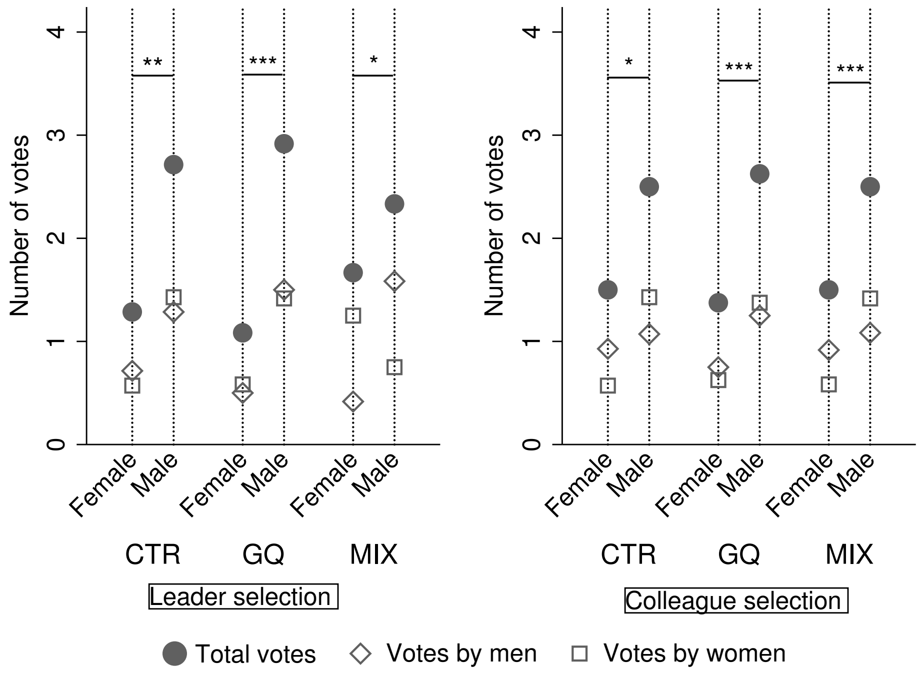

Figure B.5: **Team selection in groups with one female leader and one male leader**

Notes: The figure displays the average number of votes for leaders (left panel) and non-leaders (right panel) in stage 5 in groups with one female leader and one male leader. The brackets and stars above each line show results of Mann-Whitney U test, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

B.6 Performance in team and the cost of misselection

Regarding performance in the grid task, there are no significant differences in performance across treatments. Women perform better than men, yet the differences are not significant. Figure B.6 displays performance in grid task (stage 5) by gender and treatment. Subjects solve on average 6.77 grids correctly in CTR, 6.87 grids correctly in GQ and 7.32 grids correctly in MIX (Kruskal-Wallis test, p=0.359). Women (Men) solve 7.09 (6.46), 7.21 (6.52) and 7.76 (6.89) grids correctly in CTR, GQ, and MIX respectively (Mann-Whitney U test, p= 0.359, p= 0.240, p= 0.157 for CTR, GQ, and MIX respectively). This result confirms the false perception of competence in team selection toward women.

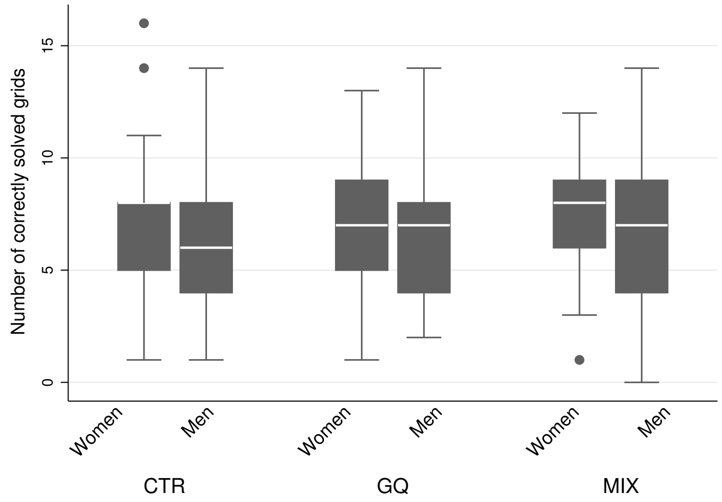

Figure B.6: **Individual performance in grid task**

Notes: The figure displays the number of correctly solved grids in stage 5 of men and women in each treatment. The upper (lower) hinges of the boxes show the 75th (25th) percentiles, the white lines inside the boxes show the median values, the upper (lower) adjacent lines show the maximum (minimum) and the dots show outliers.

To study how team (mis)selection affects team performance. We calculate the deviation in percentage points between realized team performance and optimal team performance.[26] The realized performance is team performance created by real team selection decisions, while the optimal team performance is calculated by assuming every subject selects their team optimally. In Figure B.7, we show the inefficiency due to team selection overall, and disaggregate this measure into leader selection and colleague selection.

We observe no significant differences in overall inefficiency and in inefficiency due to colleague selection across treatments.[27] Inefficiency due to leader selection is highest in CTR (25.09%), lower in GQ (15.19%), and lowest in MIX (11.30%). For both men and women, the efficiency loss because of leader selection in MIX is significantly lower than in CTR (Mann-Whitney U test, p= 0.043, p= 0.042 for women and men, respectively). The differences between GQ and CTR are not significant (Mann-Whitney U test, p= 0.286, p= 0.104 for women and men, respectively). This evidence complements Result 3 on the reduced misperception of competence against female leaders in MIX.

[26] For example, efficiency loss due to leader selection is calculated as (1 - (Performance of chosen leader/Performance of the best leader)) × 100%.

[27] The overall inefficiency is highest in CTR (17.76%), followed by GQ (15.25%) and MIX (15.24%) (Mann-Whitney U test, GQ vs. CTR: p= 0.462, GQ vs. MIX: p= 0.797, and CTR vs. MIX: p= 0.400). Inefficiency due to colleague selection is highest in MIX (16.58%), then GQ (16.05%) and CTR (13.43%) (Mann-Whitney U test, GQ vs. CTR: p= 0.552, GQ vs. MIX: p= 0.730, and CTR vs. MIX: p= 0.323).

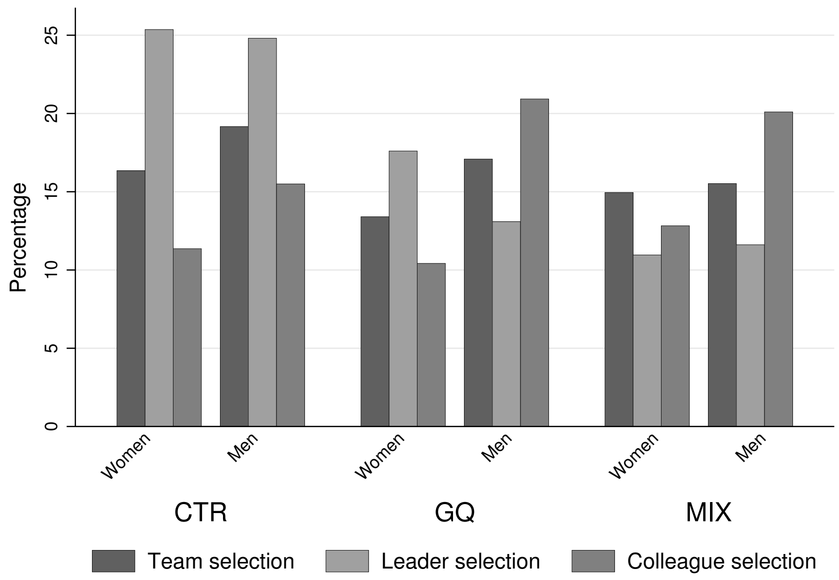

Figure B.7: **Efficiency loss due to team selection**

Notes: The figure displays the deviation in percentage points between realized team performance and optimal team performance. For example, efficiency loss due to leader selection is calculated as (1 - (Performance of chosen leader/Performance of the best leader)) × 100%.

B.7 Power analysis

In the following, we report power calculations for our main results. We follow conventional standards and fix power at 80% and alpha at 0.05. We assume two-sided test and take the mean values of the CTR treatment as the benchmark whenever possible. The power calculations were conducted after the experiment.

Efficiency: With regards to efficiency for all subjects in stage 3, we can detect differences in performance of at least 1.84 - 1.88 tasks depending on the comparison. This is about 17% of mean of CTR group (10.10). As regards winners' efficiency in stage 3, we can detect differences of at least 3.41 - 3.62 tasks depending on the comparison. This corresponds to 31%-36% of the CTR group mean for all subjects(10.10) and 23%-24% of the CTR group mean for winners (14.88). Actually observed productivity differences are small, at most 1.4% of the mean of CTR for all subjects, and at most 3.5% of the mean of CTR for winners. The data thus allows us to rule out that the different affirmative action policies have meaningful effects on efficiency.

Willingness to compete: At the aggregate level, we only have the power to detect changes in the mean of at least 21.4pp-21.7pp, for the pairwise comparison between GQ/MIX and CTR (equivalent to 39.1%-39.7% of the CTR mean of 0.55). The differences in willingness to compete between CTR and the other treatments are small throughout, 0.1pp-0.05pp in absolute value for GQ and MIX respectively, i.e., between 10% and 19% of the mean of CTR. In terms of willingness to compete of men and women separately, ex-post power analysis shows that we can detect effect sizes of at least 0.25pp-0.32pp depending on the exact pairwise treatment comparison, equivalent to 37.9%-74.3% of the mean of 0.67 for men and 0.43 for women in CTR.

Self-confidence (belief on ranking in Stage 3): At the aggregate level, we have the power to detect changes in the mean of at least 0.60-0.63 points (on a 6-point scale). This corresponds to 22.1%-23.4% of the CTR mean of 2.70. By gender, we are able to detect power of at least 0.80-0.90 points depending on the pairwise comparison. This is equivalent to 30.0%-37.2% of the CTR mean of 2.43 for men and 2.97 for women.

Team selection: With regards to leader selection at the aggregate level, we can detect effect sizes of at least 0.98 - 1.08, which is equivalent to 49% - 54% change compared to the CTR mean of 2 for the pairwise comparisons of GQ/MIX and CTR. For the pairwise comparison between men and women per treatment taking the number of votes for women as the control, we can detect effect size of at least 1.15-1.58, which is equivalent to 68.0% - 102.2% change compared to the control group means. With regards to colleague selection, we are able to detect effect sizes of at least 0.71-0.74, which is equivalent to 36%-37% change compared to the CTR mean of 2 for the pairwise comparisons of GQ/MIX and CTR. For the pairwise comparison between men and women per treatment taking the number of votes for women as the control, we can detect effect size of at least 0.84-1.04, which is equivalent to 48.2% - 69.5% change compared to the control group means.

Teamwork: For pairwise comparisons of the treatments GQ and MIX to CTR, we are able to detect effect sizes of at least 1.42-1.44, which is roughly a 21% performance change compared to the CTR mean of 6.77. Actually observed differences are 1.4% of the CTR mean for GQ, and 8.1% for MIX.

C Belief elicitation

(Translated from German, shown on screen at the end of stage 4)

In the following, we would like you to consider your performance in PART 1 (piece-rate), PART 2 (tournament), and PART 3 (choice), and guess your relative rank within your group of six. We also would like you to guess your relative rank among the three group members with the same gender/color as you.

Rank 1 is the group member with the highest performance, Rank 2 is the group member with the second highest performance, etc.

One of your guesses will be randomly chosen. If you are correct with this guess, you will also receive 1 EUR in addition to your payment from the other parts of the experiment. Thus, you should think carefully about each of the guesses.

- Which rank (from 1 to 6) within your group of six do you think you had in PART 1?
- Which rank (from 1 to 3) among the three group members with the same gender as you do you think you had in PART 1?
- Which rank (from 1 to 3) among the three group members with the same color as you do you think you had in PART 1?

Similar questions are asked about PART 2 and PART 3. The questions are programmed such that subjects can not give unreasonable answers. For example, a guessed rank of 1 in the whole group of six together with a guessed rank of 2 in group of same gender/color are not possible.

D Experimental instructions

(Translated from German)

General instructions (distributed on paper at the beginning of the experiment)

Welcome to today's experiment! Thank you for participating!

During the experiment, you and the other participants will be asked to make decisions. Your own decisions as well as the decisions of the other participants will determine your earnings, 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 your decisions and answers will remain confidential and anonymous.

The experiment consists of several parts. Additionally, you will answer a short questionnaire.

One of the parts will be selected randomly by the computer to determine your payment. Every part of the experiment is equally likely to be selected. It is therefore in your own interest to make your decisions in each part as if it was the only part.

Additionally you will receive a show-up fee of 4 EUR. This means that your total earnings will be the payment from the randomly chosen part plus the show-up fee of 4 EUR.

All other explanations will be given step-wise at the beginning of each part of the experiment. You will receive the instructions for each part in turn. You will have enough time to read the instructions carefully and to ask questions. Please do not hesitate to ask questions if something is unclear.

Please note that, as the last week, talking is not permitted. If you have questions, please do not ask them loudly but raise your hand. One of the experimenters will come to your seat to answer your question. If you do not comply with these rules you will be excluded from the experiment and you will not receive any payments.

General information regarding today's experiment

In today's experiment, your task is to add **as many five two-digit numbers as possible in a given amount of time.**

Each participant receive one color, either Green or Blue. The color will be assigned randomly to each participant. For **half of the participants**, the color will be **Green**. For **the other half**, it will be **Blue**.

We first start with a trial round. This round is for you to get used to the task. Therefore, your performance in this round is irrelevant for your final earnings. The trial round lasts 2 minutes. Afterwards, the first part of the experiment starts.

You will soon receive information about your color on the following screen.

PART 1 – Piece-rate (distributed on paper at the beginning of stage 1)

Your task in PART 1 is to add **as many five two-digit numbers as possible in five minutes.**

The use of a calculator or the similar is not allowed. You are allowed to use the provided scratch paper and pen if you like. After you have entered your answer, please click the "Confirm" button.

If PART 1 of the experiment is chosen for payment, you will receive the following payment:

$$\textbf{Payment} = \textbf{Number of correctly solved tasks} \times \textbf{0.50 EUR}$$

Your payment will not be reduced if you enter a wrong answer. We will refer to this payment as the **piece-rate payment** from now on.

You will be informed about your performance in this part at the end of the experiment.

After all questions regarding PART 1 are answered, your will start working on the task.

PART 2 – Tournament (distributed on paper at the beginning of stage 2)

As in PART 1, your task is to add as many five two-digit numbers as possible in five minutes. However, in this part your payment depends on your performance relative to the performance of other participants in your group.

Group allocation:

For the following parts of the experiment, you will be allocated to a **group with 6 members**. The groups were formed randomly and stay the same throughout the whole experiment. This means that you will form a group with the same participants for the rest of the experiment.

Each group consists of six **members**, and meets the following criteria:

- **Three** group members are women, **the other three** group members are men.
- **Three** group members are randomly assigned the color Green, **the other three** group members are randomly assigned the color Blue.

Rules of the tournament:

If PART 2 is chosen for payment, your payment depends on how high your performance is compared to the other five members of your group.

The **two** group members with **highest** performance are the two winners of the tournament.

(The content of the following part in gray differs across treatments. There is no further content for the control treatment (CTR).)

Gender quotas (GQ):

In addition, the following **special rule** is applied: **at least one winner must be a woman.**

If this is not automatically the case given the performance of the group members, then the female group member with the best performance among the three female group members will **replace** the initial second-best winner. **In this case the group**

member with the second highest performance of all six group members of your group is no longer a winner.

Mixed quotas (MIX):

In addition, the following **special rule** is applied:

- With 50% probability, rule A is applied: **at least one winner must be a woman.**
- With 50% probability, rule B is applied: **at least one winner must be a group member with the color Green.**

Rule A: If this is not automatically the case given the performance of the group members, then the female group member with the best performance among the three female group members will **replace** the initial second-best winner. **In this case the group member with the second highest performance of all six group members of your group is no longer a winner.**

Rule B: If this is not automatically the case given the performance of the group members, then the group member with the color Green and with the best performance among the three group members with the color Green will **replace** the initial second-best winner. **In this case the group member with the second highest performance of all six group members of your group is no longer a winner.**

The payment of the two winners is as follows:

$$\textbf{Payment} = \textbf{Number of correctly solved tasks} \times \textbf{1.50 EUR}$$

The **other four** members of your group get **no payment**.

If there is a tie between two group members, the winner will be determined randomly. We will refer to this payment as **tournament payment** from now on. At the end of the experiment, you will be informed about the outcome of the tournament.

In case you have any questions, please raise your hand.

PART 3 – Choice between piece-rate and tournament payment

(distributed on paper at the beginning of stage 3)

Similar to PART 1 and PART 2, your task is to add as many five two-digit numbers as possible in five minutes.

However, now you choose by yourself which payment scheme you prefer for your performance in PART 3. You can choose either the piece-rate payment (same rules as in PART 1) or the tournament payment (same rules as in PART 2).

If PART 3 is chosen for payment, your earnings will be determined as follows:

- If you choose the **piece-rate payment**, your payment is:

$$\textbf{Payment} = \textbf{Number of correctly solved tasks} \times \textbf{0.50 EUR}$$

- If you choose the **tournament payment, your earnings depend on the level of your performance in PART 3 compared to the performance of your five group members in PART 2 (tournament).** Reminder: PART 2 is the part you have just finished.

(The content of the following part in gray differs across treatments.)

Control treatment (CTR):

If your performance is higher than that of at least four other members of your group in PART 2, your payment is as follows:

Gender quotas (GQ): In addition, the following **special rule** is applied: **at least one winner must be a woman.**

If this is not automatically the case given the performance of the group members, then the female group member with the best performance among the three female group members will **replace** the initial second-best winner. **In this case the group member with the second highest performance of all six group members of your group is no longer a winner.**

If your performance in PART 3 relative to the performance of your group members in PART 2 implies you are a winner, your payment is as follows:

Mixed quotas (MIX): In addition, the following **special rule** is applied:

- With 50% probability, rule A is applied: **at least one winner must be a woman.**
- With 50% probability, rule B is applied: **at least one winner must be a group member with the color Green.**

Rule A: If this is not automatically the case given the performance of the group members, then the female group member with the best performance among the three female group members will **replace** the initial second-best winner. **In this case the group member with the second highest performance of all six group members of your group is no longer a winner.**

Rule B: If this is not automatically the case given the performance of the group members, then the group member with the color Green and with the best performance among the three group members with the color Green will **replace** the initial second-best winner. **In this case the group member with the second highest performance of all six group members of your group is no longer a winner.**

If your performance in PART 3 relative to the performance of your group members in PART 2 implies you are a winner, your payment is as follows:

$$\textbf{Payment} = \textbf{Number of correctly solved tasks} \times \textbf{1.50 EUR}$$

That means it is three times as high as the piece-rate payment.

If your performance in PART 3 relative to the performance of the other group members in PART 2 implies that **you are not a winner**, you get **no payment**.

If there is a tie between two group members, the winner will be randomly determined.

The group composition is the same as in PART 2. If you choose the tournament payment, you will be informed about the outcome of the tournament at the end of the experiment.

On the next screen, you will decide whether you choose the piece-rate payment or the tournament payment for your performance in PART 3. Then the task will begin.

In case you have any questions, please raise your hand.

PART 4 – Choice between piece-rate and tournament payment for

performance in PART 1 (distributed on paper at the beginning of stage 4)

In this part, you will not work on the task. Instead, you choose by yourself which payment scheme you prefer for your performance in PART 1. You can choose either the piece-rate payment (same rules as in PART 1) or the tournament payment (same rules as in PART 2) for your performance in PART 1.

If PART 4 is chosen for payment, your earnings will be determined as follows:

- If you choose the **piece-rate payment**, your payment is:

  Payment = Number of correctly solved tasks in PART 1 × 0.50 EUR

- If you choose the **tournament payment, your earnings depend on the level of your performance in PART 1 compared to the performance of your five group members in PART 1.**

(The content of the following part in gray differs across treatments.)

Control treatment (CTR):

If your performance in PART 1 is higher than that of at least four other members of your group in PART 1, your payment is as follows:

Gender quotas (GQ): In addition, the following **special rule** is applied: **at least one winner must be a woman.**

If this is not automatically the case given the performance of the group members, then the female group member with the best performance among the three female group members will **replace** the initial second-best winner. **In this case the group member with the second highest performance of all six group members of your group is no longer a winner.**

If your performance in PART 1 relative to the performance of your group members in PART 1 implies you are a winner, your payment is as follows:

Mixed quotas (MIX):

In addition, the following **special rule** is applied:

- With 50% probability, rule A is applied: **at least one winner must be a woman.**
- With 50% probability, rule B is applied: **at least one winner must be a group member with the color Green.**

Rule A: If this is not automatically the case given the performance of the group members, then the female group member with the best performance among the three female group members will **replace** the initial second-best winner. **In this case the group member with the second highest performance of all six group members of your group is no longer a winner.**

Rule B: If this is not automatically the case given the performance of the group members, then the group member with the color Green and with the best performance among the three group members with the color Green will **replace** the initial second-best winner. **In this case the group member with the second highest performance of all six group members of your group is no longer a winner.**

If your performance in PART 1 relative to the performance of your group members in PART 1 implies you are a winner, your payment is as follows:

Payment = Number of correctly solved tasks in PART 1 × 1.50 EUR

That means it is three times as high as the piece-rate payment.

If your performance in PART 1 relative to the performance of the other group members in PART 1 implies that **you are not a winner**, you get **no payment**.

If there is a tie between two group members, the winner will be randomly determined.

The group composition is the same as in PART 2. If you choose the tournament payment, you will be informed about the outcome of the tournament at the end of the experiment.

On the next screen, you will decide whether you choose the piece-rate payment or the tournament payment for your performance in PART 1.

In case you have any questions, please raise your hand.

PART 5 – Teamwork (distributed on paper at the beginning of stage 5)

In PART 5, you will work on a new task in which you have to solve as many counting tasks correctly as possible in five minutes, i.e. to correctly count the number of zeros ("0") in as many tables as possible. Each table consists of ten rows and ten columns, which contain either a zero ("0") or a one ("1"). Each table differs from the previous one. After you have entered your response, please click the "Confirm" button.

Leaders and colleagues:

Your group is divided into two leaders and four colleagues. The two leaders are the two winners in PART 2. The four colleagues are the four other group members who did not win in PART 2.

Team:

Each participant will form a team of three from their group of six. The team must consist of one leader and two colleagues. If you are a leader, you will choose two colleagues from your five group members for your team. If you are not a leader, you will choose one leader and one colleague from your five group members for your team.

You will not get any information about the identity of other group members and other group members will not get any information concerning your identity. The only information you will know about other group members before you make your decision is whether (s)he has won the tournament in PART 2 or not, and his (her) gender.

NOTE: Your team choice will not affect the team choice and the payment of other group members. The team choice of other group members will not affect your team choice and your payment. A leader/colleague can be chosen by more than one group member.

Payment:

Each member of your team of three will work on different tables of the counting task. Your payment in this part depends on how many correctly solved tables you and your teammates finish.

Precisely, your payment is determined as follows: you will receive 20 Euro-cent for each correctly solved table by each member of your team (including yourself). The other members of your team will also receive 20 Euro-cent for each correctly solved table that any team member (including yourself) has finished. This means each team member earns an equal amount of payment from the total correctly solved tables by the team (i.e. all three team members together).

In addition, the leaders (who won the tournament in PART 2) will each receive a bonus of 5 EUR. The colleagues will each receive a bonus of 2 EUR.

When PART 5 is chosen for payment, your payment is the sum of your bonus and your earnings (the sum of all correctly solved tables of your team × 20 Euro-Cent). At the end of the experiment, you will be informed about the overall performance of your team.

For example, if your team solve twenty tables correctly, the leader will receive the following payment:

$$\mathbf{20 \times 0.2 + 5 = 9\ EUR}$$

and the colleagues of your team will receive the following payment:

$$\mathbf{20 \times 0.2 + 2 = 6\ EUR}$$

In the following screen, you will make decision which group member are your teammates (leader and colleagues).

In case you have any questions, please raise your hand.

E Questionnaire

The questionnaire at the end of the experiment contains the following items (translated from German):

1. **Fairness perception:** In all treatments, we first elicit fairness perception of the competition rule in use, then of the other two competition rules, using the following scale:

   completely unfair ☐ ☐ ☐ ☐ ☐ ☐ ☐ *completely fair*

2. **Risk preference, general risk question:** same wording as in German Socio-Economic Panel questionnaire (SOEP, see, for example, Wagner, Frick, and Schupp (2007))
   How do you evaluate yourself? Are you generally a risk-seeking person or do you try to avoid risks? The leftmost box means "not at all risk-seeking" and the rightmost "very risk-seeking". With the boxes in between, you can graduate your statement.

   not at all risk-seeking ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ *very risk-seeking*

3. **Risk preference, incentivized choice list:** Subjects make eleven, pairwise decisions between a lottery with a fifty-fifty chance of winning either 2 EUR or 7 EUR and a safe payment. The safe payment increases in 0.5 EUR increments, ranging from 2 EUR to 7 EUR.

4. **Cognitive reflection test:** see Frederick (2005).

   After subjects finish the test, we elicit their belief about individual rank compared to the other participants in the experiment, with three possible answers: 0-33.3%, 33.4-66.6%, and 66.7-100% (implying they answer one/two/three out of three questions in the test correctly, respectively.)

5. **Social preference** (survey question, Falk et al., 2018)
   Question 1: Imagine the following situation: Today you unexpectedly received 1000 EUR. How much of this amount would you donate to a good cause? (Values between 0 and 1000 are allowed).
   Question 2: Please think about what you would do in the following situation. You are in an area you are not familiar with, and you realize that you lost your way. You ask a stranger for directions. The stranger offers to take you to your destination. Helping you costs the stranger about 20 EUR in total. However, the stranger says he or she does not want any money from you. You have six presents with you. The cheapest present costs 5 EUR, the most expensive one costs 30 EUR. Do you give one of the presents to the stranger as a "thank you" gift?
   Which present do you give to the stranger?
   1. No, would not give present
   2. The present worth 5 EUR
   3. The present worth 10 EUR
   4. The present worth 15 EUR
   5. The present worth 20 EUR
   6. The present worth 25 EUR
   7. The present worth 30 EUR

6. **Socio-demographics:** age, final grade point average at academic high school, last math grade at academic high school, field of study, monthly disposable amount of money, political orientation, number of experiments already participated in the same lab.

7. **Positive and negative reciprocity, trust, and belief on own mathematical ability** (survey question, Falk et al., 2018)

PREVIOUS DISCUSSION PAPERS

405 Trieu, Chi, Who's Who: How Uncertainty About the Favored Group Affects Outcomes of Affirmative Action, August 2023.
Forthcoming in: The Journal of the Economic Science Association.

404 Fischer, Kai, Martin, Simon and Schmidt-Dengler, Philipp, The Heterogeneous Effects of Entry on Prices, July 2023.

403 García-Vega, María, Gupta, Apoorva and Kneller, Richard, Is Acquisition-FDI During an Economic Crisis Detrimental for Domestic Innovation?, July 2023.

402 Haucap, Justus and Stiebale, Joel, Non-price Effects of Mergers and Acquisitions, July 2023.

401 Simion, Ștefania and Sulka, Tomasz, Multidimensional Cognitive Ability, Intermediate Channels, and Financial Outcomes, May 2023.

400 Bartling, Björn, Cappelen, Alexander W., Hermes, Henning, Skivenes, Marit and Tungodden, Bertil, Free to Fail? Paternalistic Preferences in the United States, May 2023.

399 Kandelhardt, Johannes, Flexible Estimation of Random Coefficient Logit Models of Differentiated Product Demand, May 2023.

398 Hermes, Henning, Lergetporer, Philipp, Mierisch, Fabian, Peter, Frauke and Wiederhold, Simon, Discrimination on the Child Care Market: A Nationwide Field Experiment, April 2023.

397 Schmal, W. Benedikt, Haucap, Justus and Knoke, Leon, The Role of Gender and Coauthors in Academic Publication Behavior, March 2023.

396 Magin, Jana Anjali, Neyer, Ulrike and Stempel, Daniel, The Macroeconomic Effects of Different CBDC Regimes in an Economy with a Heterogeneous Household Sector, March 2023.

395 Dertwinkel-Kalt, Markus and Wey, Christian, Resale Price Maintenance in a Successive Monopoly Model, February 2023.
Forthcoming in: Journal of Industrial Economics.

394 Hermes, Henning, Krauß, Marina, Lergetporer, Philipp, Peter, Frauke and Wiederhold, Simon, Early Child Care and Labor Supply of Lower-SES Mothers: A Randomized Controlled Trial, December 2022.

393 Chowdbury, Shyamal, Schildberg-Hörisch, Hannah, Schneider, Sebastian O., and Sutter, Matthias, Information Provision Over the Phone Saves Lives: An RCT to Contain COVID-19 in Rural Bangladesh at the Pandemic's Onset, November 2022.

392 Normann, Hans-Theo and Sternberg, Martin, Human-Algorithm Interaction: Algorithmic Pricing in Hybrid Laboratory Markets, October 2022.
Published in: European Economic Review, 152 (2023), 104347.

391 Hunold, Matthias and Petrishcheva, Vasilisa, Foreclosure and Tunneling with Partial Vertical Ownership, September 2022.

390 Haucap, Justus and Heldman, Christina, The Sociology of Cartels, August 2022.

389 Döpper, Hendrik, Sapi, Geza and Wey, Christian, A Bargaining Perspective on Vertical Integration, May 2022.
Forthcoming in: Canadian Journal of Economics.

388 Bachmann, Ronald, Gonschor, Myrielle, Lewandowski, Piotr and Madoń, Karol, The Impact of Robots on Labour Market Transitions in Europe, May 2022.

387 Fremerey, Melinda, Hörnig, Lukas and Schaffner, Sandra, Becoming Neighbors with Refugees and Voting for the Far-Right? The Impact of Refugee Inflows at the Small-Scale Level, April 2022.

386 Fischer, Kai, Alcohol Prohibition and Pricing at the Pump, March 2022.

385 Cobb-Clark, Deborah A., Dahmann, Sarah C., Kamhöfer, Daniel A. and Schildberg-Hörisch, Hannah, The Determinants of Population Self-Control, March 2022.

384 Sulka, Tomasz, Planning and Saving for Retirement, March 2022.

383 Cattan, Sarah, Kamhöfer, Daniel A., Karlsson, Martin and Nilsson, Therese, The Long-term Effects of Student Absence: Evidence from Sweden, March 2022.
Published in: Economic Journal, 133 (2023), pp. 888-903.

382 Martin, Simon and Rasch, Alexander, Collusion by Algorithm: The Role of Unobserved Actions, March 2022.

381 Haucap, Justus, Nedic, Radivoje and Şimşek, Talha, An Empirical Analysis of German Casino Locations, March 2022.
Published in: European Journal of Law and Economics, 55 (2023) pp. 291-311.

380 Haucap, Justus, Heldman, Christina and Rau, Holger A., Gender and Collusion, March 2022.

379 Schain, Jan Philip, Foreign Institutional Investors and the Great Productivity Slowdown, November 2022 (First Version February 2022).

378 Neyer, Ulrike and Stempel, Daniel, How Should Central Banks React to Household Inflation Heterogeneity?, January 2022.

377 Döpper, Hendrik and Rasch, Alexander, Combinable Products, Price Discrimination, and Collusion, January 2022.

376 Cobb-Clark, Deborah A., Dahmann, Sarah C., Kamhöfer, Daniel A. and Schildberg-Hörisch, Hannah, Self-Control and Unhealthy Body Weight: The Role of Impulsivity and Restraint, January 2022.
Forthcoming in: Economics & Human Biology.

375 Martin, Simon and Schmal, W. Benedikt, Collusive Compensation Schemes Aided by Algorithms, December 2021.

374 Kouli, Yaman and König, Jörg, Measuring European Economic Integration 1880 - 1913 – A New Approach, December 2021.

373 Hunold, Matthias and Schad, Jannika, Single Monopoly Profits, Vertical Mergers, and Downstream Entry Deterrence, December 2021.

372 Werner, Tobias, Algorithmic and Human Collusion, November 2021.

371 Jovanovic, Dragan, Wey, Christian and Zhang, Mengxi, On the Social Welfare Effects of Runner-up Mergers in Concentrated Markets, September 2021.
Published in: Economics Bulletin, (2021), Article 201.

370 Jeschonneck, Malte, Collusion among Autonomous Pricing Algorithms Utilizing Function Approximation Methods, August 2021.

369 Gösser, Niklas, Gürer, Kaan, Haucap, Justus, Meyring, Bernd, Michailidou, Asimina, Schallbruch, Martin, Seeliger, Daniela and Thorwarth, Susanne, Total Consumer Time – A New Approach to Identifying Digital Gatekeepers, August 2021.

368 Fischer, Kai, Reade, J. James and Schmal, W. Benedikt, The Long Shadow of an Infection: COVID-19 and Performance at Work, August 2021.
Published in: Labour Economics, 79 (2022), 102281.

367 Suedekum, Jens, Place-Based Policies – How to Do Them and Why, August 2021.

366 Heiss, Florian, Ornaghi, Carmine and Tonin, Mirco, Inattention vs Switching Costs: An Analysis of Consumers' Inaction in Choosing a Water Tariff, July 2021.

365 Cobb-Clark, Deborah A., Dahmann, Sarah C., Kamhöfer, Daniel A. and Schildberg-Hörisch, Hannah, Sophistication about Self-Control, July 2021.

364 Bie, Xiaodong and Ciani, Andrea, Born Similar, Develop Apart: Evidence on Chinese Hybrid Exporters, July 2021.

363 Ali, Nesma and Stiebale, Joel, Foreign Direct Investment, Prices and Efficiency: Evidence from India, July 2021.

362 Banerjee, Ritwik, Ibanez, Marcela, Riener, Gerhard and Sahoo, Soham, Affirmative Action and Application Strategies: Evidence from Field Experiments in Columbia, April 2021.

361 Wellmann, Nicolas and Czarnowske, Daniel, What Would Households Pay for a Reduction of Automobile Traffic? Evidence From Nine German Cities, March 2021.

360 Haucap, Justus, Moshgbar, Nima and Schmal, Wolfgang Benedikt, The Impact of the German "DEAL" on Competition in the Academic Publishing Market, March 2021.
Published in: Managerial and Decision Economics, 42 (2021), pp. 2027-2049.

359 Korff, Alex, Competition in the Fast Lane – The Price Structure of Homogeneous Retail Gasoline Stations, January 2021.

Older discussion papers can be found online at:
http://ideas.repec.org/s/zbw/dicedp.html

Heinrich-Heine-Universität Düsseldorf

Düsseldorfer Institut für
Wettbewerbsökonomie (DICE)

Universitätsstraße 1, 40225 Düsseldorf

ISSN 2190-992X (online)
ISBN 978-3-86304-404-6

www.dice.hhu.de