

NBER WORKING PAPER SERIES

LEANING IN OR NOT LEANING OUT? OPT-OUT CHOICE FRAMING ATTENUATES
GENDER DIFFERENCES IN THE DECISION TO COMPETE

Joyce He
Sonia Kang
Nicola Lacetera

Working Paper 26484
<http://www.nber.org/papers/w26484>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2019

We are grateful to the the Institute on Gender and the Economy (GATE) and Center for Behavioural Economics (BEAR) at at the Rotman School fo the University of Toronto for their financial support. We are grateful to Sarah Kaplan, Mario Macis, Justin Sydnor, and participants to seminars at the Max Plank Institute in Munich (Germany) for their comments. Study 2 was registered in the AEA-RCT registry with the number AEARCTR-0004030. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2019 by Joyce He, Sonia Kang, and Nicola Lacetera. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Leaning In or Not Leaning Out? Opt-Out Choice Framing Attenuates Gender Differences
in the Decision to Compete

Joyce He, Sonia Kang, and Nicola Lacetera

NBER Working Paper No. 26484

November 2019

JEL No. C91,D03,D91,J16,J24,J82,M5

ABSTRACT

In most organizations, promotions often require self-nomination and competition among applicants. However, research on gender differences in preferences for competition suggests that this process might result in fewer women choosing to participate. We study whether changing promotion schemes from a default where applicants must opt in (i.e., self-nominate) to a default where applicants must opt out (i.e., they are automatically considered for promotion, but can choose not to be considered) attenuates gender differences. In our first experiment, although women are less likely than men to choose competitive environments under the traditional opt-in framing, in the opt-out system both women and men have the same participation rate as men in the opt-in system. The increase in participation of women into competition is not associated with negative consequences on performance or well-being. In our second experiment, we show that opt-out framing does not entail penalties from evaluators making decisions about whom to hire. These results support the promise of choice architecture to reduce disparities in organizations. More generally, our findings suggest that gender differences in attitudes toward completion may be context-dependent.

Joyce He
University of Toronto - Rotman School
105 St. George Street
Toronto, ON
Canada
joyce.he@mail.utoronto.ca

Sonia Kang
University of Toronto
Institute for Management and Innovation
3359 Mississauga Road
Mississauga, ON
Canada
sonia.kang@rotman.utoronto.ca

Nicola Lacetera
University of Toronto
Institute for Management and Innovation
3359 Mississauga Road, Room KN 235
Mississauga, ON L5L 1C6
CANADA
and NBER
nicola.lacetera@utoronto.ca

1. Introduction

Only three women have ever received the Nobel Prize in Physics, the most recent being Donna Strickland. Strikingly, when she was awarded the Prize, Professor Strickland held the rank of Associate Professor. The positive reaction to her win came with puzzlement over how such an eminent scientist was not yet a *Full* Professor. Responding to a question from a BBC reporter about this, Professor Strickland explained simply: “I never applied” (BBC World Service 2018). As with promotion from Associate to Full Professor in many universities, most types of promotions require self-nomination via an application. However, past research shows that this process could disadvantage women. For example, compared to men, women have a lower propensity to compete and take risk (Flory et al. 2015, Niederle and Vesterlund 2007, Samek 2017), are less (over)confident (Niederle and Vesterlund 2007, Reuben et al. 2012), and less likely to self-promote or exaggerate accomplishments (Daubman et al. 1992, Moss-Racusin and Rudman 2010). These findings suggest that women might be less inclined to apply for promotions and competitive selection processes because of the typical requirement to self-nominate, promote, and compete, holding constant performance and ability (Bosquet et al. 2018, Exley et al. 2016).

In the absence of overall gender differences in ability and performance, the lack of women choosing to apply for competitive selection processes (i.e., promotions) is costly because organizations are unable to tap into the full pool of talent. For instance, if high-performing women do not apply for competitive processes, organizations cannot select from the full range of high-performing employees in advancement pools. Indeed, research finds that affirmative action (i.e., quotas) can boost the supply of high-performers because it corrects for lack of minority candidates who would otherwise not have self-selected into the competition (Niederle et al. 2013). Furthermore, gender differences in the choice to compete result in a reduction in diversity and its associated benefits (Bernardi et al. 2006, Dezsö and Ross 2012, Weber and Zulehner 2010). More broadly, on a societal level, the gender gap in competition contributes to and perpetuates occupational segregation by gender, both horizontally (across fields that differ in competitiveness) and vertically (within more senior management positions) (Flory et al. 2015, Samek 2017).

To date, interventions to reduce gender disparities in hiring, evaluation, and promotion have focused mainly on diversity programs, usually consisting of diversity or unconscious bias training, and have failed to show consistent positive results (Dobbin and Kalev 2016). In fact, diversity training can result in a host of unintended negative consequence, such as reduced

diversity, negative attitudes and behaviors towards minorities for which these diversity programs are intended to benefit, and increase activation of stereotypes that are meant to be de-activated or suppressed (Kalev et al. 2006, Leslie 2018, Macrae et al. 1994, Sanchez and Medkik 2004). Further, even when diversity training is effective in shifting some attitudes, it often does not translate into meaningful actions and behaviors (Chang et al. 2019).

Another popular approach is training women to be more active and visible at work, or to “lean in” (Sandberg 2013). Although this approach focuses on empowering women to negotiate better, talk more in meetings, and be more assertive, existing theory and corroborating evidence suggests that women often face backlash for adopting these more stereotypically male characteristics and behaviors (Moss-Racusin and Rudman 2010, Rudman 1998, Rudman and Glick 2001). Further, lean-in messages attribute responsibility for both causing and fixing the problem on women themselves (Fitzsimons et al. 2018, Kim et al. 2018).

In this work, we depart from approaches that attempt to “fix the women” or, more generally, change people’s minds; we focus instead on structural changes to the framing of the decision to compete for a promotion (Bohnet 2016, Chugh 2018). We apply well-established findings in behavioral science on choice architecture, and in particular the finding that behavior is strongly affected by defaults (Samuelson and Zeckhauser 1988, Thaler and Sunstein 2008), to the choice of applying for a promotion.

For promotion and the choice to enter a competition, the most typical choice architecture is an opt-in choice: individuals must actively apply for the promotion, and the default choice is to not to apply. We explore whether using an opt-out framing (i.e., making competition the default) leads to greater participation in competition (Downs et al. 2009, Johnson et al. 2002). Opt-out framing builds on the idea of a default bias by encouraging enrollment into a desired choice by changing the default to that of participation. This choice architecture has been highly successful, for example, in increasing enrollment in retirement savings and organ donation programs (Choi et al. 2002, Johnson and Goldstein 2003). Importantly, this intervention tweaks the framing of the decision, while maintaining each individual’s ability to choose.

We test the hypothesis that a choice architecture where applicants must opt-out of competition will change gender differences in the decision to compete by introducing one fundamental change to the choice framing of the well-established experimental paradigm developed by Niederle and Vesterlund (2007). This paradigm has consistently identified a gender

difference in preference for competition (Healy and Pate 2011, Niederle et al. 2013). We examine decisions about compensation for performance in a competitive framework because relative performance evaluations and promotions have the same underlying “tournament” mechanism (Lazear and Rosen 1981), whereby multiple individuals “compete” and are compared for one single position.

In our first experiment, participants ($n = 482$) completed three stages of a timed math task. In stage 1, participants received \$0.50 per correct answer in a non-competitive task with piece rate compensation. The second stage used competitive tournament compensation; we compared the focal participants’ score to those of three other participants in the room; if the participant’s score was highest, they won the tournament and received \$2.00 per correct answer; if not, they received \$0. Finally, before stage 3 (and without knowing their rank in stage 2’s tournament), participants were asked to choose how they wanted to be compensated—via non-competitive piece rate compensation or via competitive tournament compensation. Here, we randomly assigned participants to one of two choice architecture conditions: 1) Participants were automatically enrolled into non-competitive piece-rate compensation, but had the option to opt-in to competitive compensation; 2) Participants were by default in a competitive tournament-based compensation scheme, but could opt out to non-competitive piece-rate pay. Following this choice, participants completed stage 3 of the math task. The choice to compete in stage 3 is our main outcome of interest. After stage 3, participants guessed their relative rank in the stage 2 tournament compensation scheme, and filled out a brief anxiety scale.

Under the opt-in choice frame, we replicate previous research to show that women were much less likely than men to select the competitive tournament compensation. However, under the opt-out frame, both men and women had roughly the same participation rate in the competition as men in the opt-in condition, thus eliminating the gender difference in decision to compete. We further find that nudging women to compete more in the opt-out frame did not negatively affect their performance or their broader well-being. In a second experiment, we examined whether the choice architecture of the decision to compete affected the likelihood that evaluators would choose to promote a woman, and found no such difference between an opt-in and an opt-out choice architecture.

Our findings have implications for the design of promotion schemes in organizations. Specifically, an opt-out promotion scheme may increase the number of women who apply for a

promotion; as such, it would allow organizations to take advantage of a broader talent pool. Moreover, because this intervention focuses on removing biases built into the system rather than targeting a specific group, these results likely have positive spillover effects, such as increased entry of any other underrepresented group that similarly have a lower preference for competition, and perceptions of fairness, transparency and justice (Leslie 2018).

More broadly, our findings imply that the lower propensity of women to compete may not be absolute, but, instead, contextual, and in particular it may depend on how organizations define the choice architecture around the decision to compete.

In Section 2, we describe the design and results of our first experiment where we examine the effect of choice architecture on applicants' decisions to enter into competitive tournaments. Section 3 describes our second experiment in which we examine how choice architecture affects evaluators' decisions. In Section 4, we discuss the implications of our findings.

2. Study 1: Gender, preferences for competition, and choice frame

In our first experiment, we directly examine whether choice architecture affects the decision to compete by men and women. Our design follows the paradigm of Niederle and Vesterlund (2007) with the key addition of the random assignment of participants to an opt-in choice framework (as in Niederle and Vesterlund) versus an opt-out framework.¹

2.1 Experimental design

Participants. Participants were 482 undergraduate students from a large Canadian university (55.4% women; $M_{\text{age}} = 19$, $SD = 1.64$; ethnicity: 67.4% Asian, 19.9% Caucasian, 1.45% Hispanic/Latino, 1.24% African American, 9.96% indicated "other"). Participants received one course credit for participating and earned financial compensation from one randomly selected stage of the task.

Procedure. We used zTree (version 4.0) to program the experiment (Fischbacher 2007), largely following the paradigm in Niederle and Versterlund (2007), which examined gender differences in overconfidence and preference for competition. The experimental task consisted of adding five

¹ Detailed instructions to the participants of both Study 1 and Study 2 are in the Appendix.

two-digit numbers. Participants could not use a calculator, but we provided them with scrap paper; they had up to 5 minutes to complete as many questions as they could. At the end of each stage, participants saw their own final score (i.e., the number of correct answers) for that stage. In each stage, participants saw only their own absolute score and did not see their relative performance until the end of the experiment. Participants completed four stages of this same task. Following Niederle and Vesterlund (2007) as well as many other similar experiments, participants were informed that their monetary rewards would depend on their performance in one of the four stages, randomly determined; this would ensure that participants had high-powered incentives in all stages. The compensation scheme, however, was different for each stage as described below.

Stage 1 – Piece Rate: If stage 1 was randomly selected for compensation, participants would receive \$0.50 for each correct answer.

Stage 2 – Tournament: If stage 2 was selected for payment, a participant's score for that stage would be compared to three other randomly chosen competitors' scores. If the participant had the highest score in that stage compared to the other three competitors, they would receive \$2 per correct answer. If they did not hold the highest score, they would receive \$0. In the case of a tie, the winner was determined randomly among the top scorers. Note that the expected payoffs for the tournament and piece-rate compensation are equal. Moreover, the participants for a given group of four were randomly selected with replacement. Therefore, a given individual could be in the comparison group for more than one focal participant.

Stage 3 – Choice: Before proceeding to the task, participants were asked to choose their compensation scheme for the addition exercise. In the original design of the experiment (Niederle and Vesterlund 2007), participants were told that they could choose piece-rate (50 cents per correct answer) or tournament (\$2 per correct answer if the focal participant's score exceeded that of the other group members in the *stage two* tournament; winners were chosen randomly in the case of a tie). Here, we administered our key manipulation, i.e. whether the choice to enter the competitive environment (the tournament payment scheme) was expressed using opt-in or opt-out framing. We randomly assigned the participants to either the opt-in or the opt-out condition.

Opt-In Framing Condition. In the opt-in framing condition, participants were told that by default, if stage 3 was randomly selected for payment, they would receive \$0.50 per correct answer. In other words, the default was the non-competitive, piece-rate compensation. Participants could choose instead to opt in to the competitive, tournament compensation scheme. Further, they were told that if they chose to compete, their stage 3 performance would be compared with the stage 2 performance of the three other participants. We chose to compare against stage 2 performance to avoid instances where not all competitors chose to compete for stage 3. Niederle and Vesterlund (Niederle and Vesterlund 2007) also note that by comparing the performance of a focal participant to the correct answers of three participants in the previous stage, one can identify preferences for being compared separately from preferences for direct competition. If a participant wanted a tournament-based compensation, they had to check a box to indicate this. Otherwise, they just had to press the “next” button to proceed to the next page.

Opt-Out Framing Condition. In the opt-out framing condition, participants were told that by default, if stage 3 was randomly selected for payment, their performance would be compared to the same three participants’ from the previous stage, and if they received the highest score they would receive \$2 per correct answer and \$0 if they did not receive the highest score. In other words, the default was the competitive, tournament compensation. Participants could choose to opt out of the tournament and return to the non-competitive, piece-rate compensation scheme. If they preferred to opt out of the tournament compensation scheme, again they had to check a box to press the “next” button to proceed.

Stage 4 – Choice: In the final stage of the experiment, participants could re-submit their stage 1 performance for compensation. They could choose to submit their stage 1 performance to either the piece-rate compensation or a tournament compensation, where their performance would be compared to three other participants’ stage 1 performance. Note therefore that there was not an actual task in stage 4.

After the four stages, we asked participants to guess their rank compared to the other competitors. With this information, we could compare their guessed rank to their actual rank to

obtain a measure of (over)confidence. Participants guessed their rank (1=best, 2=second best, 3=third best, or 4=fourth best) in stage 1 as well as in stage 2.

Finally, we measured the well-being of the participants in terms of their perceived anxiety during the experiment. We relied on a six-item version of the State Anxiety Inventory (SAI) (Marteau and Bekker 1992, Spielberger and Gorsuch 1983). The six-item version of the SAI is highly correlated with the full version of the scale, and has high internal consistency (alphas above .90) (Marteau and Bekker 1992, Tluczek et al. 2009). We asked participants to read the statements and indicate how they felt during the experiment, on a scale from 1 (*not at all*) to 4 (*very much*). Sample items include “During the experiment, I felt calm” (reverse scored), “During the experiment, I was tense”, and “During the experiment, I felt upset”. In our sample, the six-item scale had an internal reliability of $\alpha = .68$. Upon closer examination, it appeared that the item “I felt content” had poor item-total correlation compared to the rest of the other items. Thus, we removed it from the scale (Johnson et al. 2008). The five-item scale had good reliability ($\alpha = .73$).

2.2 Results

Figure 1 shows the cumulative distribution of correct answers by gender in stage 1 (where all participants received piece-rate compensation), and in stage 2 (where compensation was tournament-based for all). Although men’s average performance ($M = 8.86$, $SD = 3.44$) in stage 1 was slightly higher than women’s ($M = 8.24$, $SD = 3.67$) (p from two-tailed t-test = 0.06), there was no significant difference in the overall distribution (p from Kolmogorov-Smirnov test = 0.35). In stage 2, both men and women attained higher average performance ($M = 10.24$, $SD = 3.75$) than under piece rate ($M = 8.52$, $SD = 3.58$, p from paired two-tailed t-test < 0.001). However, there were again no significant differences in score between men ($M = 10.43$, $SD = 3.80$) and women ($M = 10.09$, $SD = 3.70$; p from two-tailed t-test = 0.33). This evidence is consistent with men and women performing similarly in the task.

Figure 2 reports the key outcome of our study: the share of men and women who chose a tournament-based compensation scheme in stage 3, by experimental condition. The findings from the condition where participants had to opt-in to compete are similar to the evidence from previous experiments using the same paradigm, with far fewer women than men choosing tournament-based payment: 46.72% vs. 72.48% (p from two-tailed t-test < 0.001).² In contrast, the proportions of

² In the experiment of Niederle and Vesterlund (2007), 73% of men selected the tournament, against 35% of women.

women and men choosing tournament compensation in the opt-out condition were statistically indistinguishable, and similar to the percentage of men choosing tournament in the basic opt-in scheme: 75.38% of women vs. 76.42% of men (p from two-tailed t -test = 0.85).

The probit regression estimates in Table 1 confirm the descriptive evidence of Figure 2; controlling for performance in stages 1 and 2, as well as for the guessed rank in stage 2 (see column 4 in particular), we estimate women to be 24.7 percentage points less likely to participate in the tournament than men in the opt-in condition (p from two-tailed z -test < 0.001), whereas there is no significant gender difference in participation in the opt-out condition (p from two-tailed z -test = 0.85). Note also that participants who predicted better performance in stage 2 (lower guessed rank) were significantly more likely to choose a tournament in stage 3.^{3,4}

One concern with choice architecture manipulations is that they might lead individuals to make choices that negatively affect their performance or broader well-being (Mols et al. 2015, Selinger and Whyte 2011). Our evidence shows no such consequences from introducing an opt-out frame. First, Figure 3 reports participants' average performance in stage 3 by condition, gender, and compensation choice. Consistent with findings for stages 1 and 2, performance was higher for participants choosing tournament than piece rate compensation in stage 3; however, there were no significant differences by gender or experimental condition. In other words, women who were “nudged” to compete in the opt-out condition did not perform worse because of our treatment.⁵

Second, we examined the possibility that choosing a tournament in the opt-out condition led more women to make a choice that was not payoff-maximizing for them, in light of recent research finding that men tend to engage in suboptimal entry into negotiations by engaging too often, whereas women avoid negotiations but do so correctly (Exley et al. 2016). We established the payoff-maximizing choice of the participants as follows (see also Niederle and Vesterlund 2007). We ran a probit regression of whether a participant won their tournament in Stage 2 on the number of correct responses in that stage, to predict the likelihood of winning a tournament in Stage 3 given their performance in Stage 3 (recall that individuals who chose the tournament in

³ Appendix Table A2 shows that women predicted a lower ranking in stage 2, and were slightly less overconfident, controlling for absolute performance.

⁴ Figure A2 in the Appendix show that the probability that a tournament winner in stage 3 was a woman was substantially higher in the opt-out than in the opt-in condition. The share of female participants who chose a tournament and won was similar in the two conditions.

⁵ Appendix Table A1 reports the regression estimates. Figure A1 reports additional evidence of the absence of underlying gender differences in ability, from a simulation exercise.

Stage 3 had their performance compared against three participants from Stage 2). We then calculated the expected payoff from choosing a tournament or a piece-rate compensation for each participant. For the piece rate, the expected payoff was \$0.50 *times* the number of correct responses in Stage 3. The expected payoff from a tournament was \$2 *times* the number of correct responses in Stage 3 *times* the predicted probability of winning the tournament for that participant in Stage 3. We classified a participant as having made their payoff-maximizing choice if they selected the compensation scheme that gave them the higher expected payoff.⁶ Table 2 shows that in the opt-out condition, a higher proportion of women than men who chose tournament-based compensation in stage 3 made a payoff-maximizing choice as compared to what they would have gained from piece-rate compensation for the same number of correct answers. Thus, we do not find evidence that choosing a tournament in the opt-out condition led women to make choices that were not payoff-maximizing.

Third, compared to the counterfactual choices, participants earned more with the compensation method that they selected, particularly when selecting tournament. For the participants who selected a tournament-based compensation scheme, the “counterfactual” payoff is the number of correct responses that they gave in Stage 3 multiplied by \$0.5. For the participants who selected piece rate, the counterfactual payoff is zero if they would have not won the tournament in the group to which they would be assigned, and equal to the number of their correct answers multiplied by \$2 had they been the winners of the groups to which they were assigned. Figure 4 and Table 3 show the results of this analysis. Women in the opt-out condition achieved a higher additional surplus than men did; the overall additional monetary surplus was slightly greater, and more equally distributed between men and women, in the opt-out condition.

Finally, we use the anxiety measures that we collected at the end of the experiment as a measure of well-being. Figure 5 shows participants’ average reported anxiety at the end of the experiment, again by gender and condition. The graphs show higher average levels of anxiety for women than for men (overall averages 1.73 and 1.56, respectively; p from two-tailed t-test =

⁶ In cases where the expected payoffs from the two compensation scheme were close to each other (less than \$1 in absolute difference; this happened if a participant solved 12 questions correctly), we randomized the assignment to having made the payoff maximizing decision or not, with ex-ante 50% probability (we interpreted small differences as making a participant indifferent between the two schemes).

0.008); however, there was no significant average difference by condition or choice of compensation scheme in stage 3.⁷

We also further explored heterogeneity in the choice of a compensation scheme by gender and condition, by interacting the indicators for the gender-condition pairs with the performance (number of correct answers) in stages 1 and 2, as well as the guessed rank in stage 2. The regression estimates in Table 4 show that the absence of a gender difference in choosing a tournament in stage 3 under the opt-out frame derives especially from women with higher positive difference in performance between stages 2 and 1 (i.e., women who had improved in performance from stage 1 to stage 2), and those who guessed a worse rank for their performance in the stage 2 tournament.⁸

2.3 Discussion

Our results support previous evidence that even when women and men perform equally, women are less likely to select into competition. However, the aversion to competitive performance pay is contingent on how the competition choice is presented: an opt-out framing, whereby the default is to participate in the competitive scheme, eliminates this commonly observed gender difference. Critiques of the “nudging” approach have argued that nudges might be associated with negative consequences for behavior and well-being, but our additional analyses demonstrate that our intervention did not have such effects. These results resemble those found in Niederle, Segal and Vesterlund (2013), where they find that guaranteeing women equal representation among winners (i.e., affirmative action) can increase the number of women who choose to apply. Rather than introducing policy changes that specifically target women and establish a quota, however, we find the same increase in women’s participation with a change in the choice architecture for competitive compensation schemes.

Although we did not see unintended negative consequences for women’s behavior and well-being, one further concern arises in the potential for negative downstream consequences on

⁷ Regression estimates are in Appendix Table A1.

⁸ Although the compensation choices in stage 4 were not a relevant outcome of interest for this study, we report regression estimates of that choice as a function of gender, condition, and choice made in stage 3. Recall that the choice in stage 4 was whether to switch the compensation mode for the performance in stage 1 to a tournament, therefore this was an opt-in frame for all participants. Appendix Table A3 shows that women were still less likely than men to choose a tournament-based compensation, but by a smaller difference than in the opt-in condition in stage three. This is consistent with only a partial “carry over” effect of stage 3 (at least for the women in the opt-out condition in that stage). Women who chose the tournament in stage 3 were more likely to choose a tournament scheme in stage 4 than those who chose piece rate in stage 3, but the difference in propensity between these two groups was smaller than for men.

evaluators' perceptions and decisions about whom to promote. In particular, an opt-out promotion scheme might lead evaluators to attribute lower motivation and desire for competition and promotions, and this may result in a lower likelihood of promoting women compared to an opt-in promotion scheme. We address these issues in our second study.

3. Study 2: Does Competition Choice Architecture Negatively Impact Evaluation of Women?

To rule out the possibility that an opt-out choice framework might be associated with negative evaluations, we designed a second study using a yoked experimental design (Anicich et al. 2015, DeCelles et al. 2019, Harmatz and Lapuc 1968) to examine how opt-out framing impacts promotion decisions made by evaluators.

3.1 Experimental Design

Participants. We recruited 1,007 participants from Amazon's Mechanical Turk (MTurk) to participate in our study. Research has shown that participants from MTurk are a diverse and high-quality sample when compensated adequately and comprehension and attention checks are in place (Buhrmester et al. 2011). In adherence to our preregistration,⁹ we excluded all participants who failed our comprehension check, because our manipulation of the choice architecture of the competition was contingent on understanding the instructions. One hundred and eighty participants were excluded, resulting in a final sample of 827 participants (42% women; $M_{age} = 37$, $SD = 11$; primarily Caucasian (79%)). Participants received \$1.50 for participating in the 15-minute survey, and were eligible to earn a bonus of \$0.75 based on their choices in the task.

Procedure. We adapted a 2-stage design used in previous research on gender and selection decision (Bohnet et al. 2016), and programmed the experiment in oTree (Chen et al. 2016). Specifically, we assigned participants in Study 1 to the role of applicants (henceforth we refer to them as "applicants"), whereas participants of Study 2 assumed the role of evaluators (henceforth referred to as "evaluators"). Specifically, we asked evaluators to make a choice to promote an applicant to a higher position with a higher wage. We presented them brief profiles of the

⁹ This experiment was preregistered at the American Economic Association's Trial Registry (number 0004030).

applicants from Study 1, which informed them of applicants' past absolute performance on the math task in Stage 2 of Study 1,¹⁰ and information on the applicants' age, gender, year of study, and major field of study. Their central task was to make a selection decision on whom to promote. If an evaluator selected an applicant whose Stage 3 performance in Study 1 was higher than the other applicants' Stage 3 performance in their group, then they would be eligible to receive a bonus of \$0.75.

Evaluators were shown three groups of four applicants, and asked to make three promotion decisions. The \$0.75 bonus compensation would be based on one of the three decisions, chosen at random (evaluators would only be eligible for the bonus if they selected a tournament winner in at least one of the choices). This would make all three choices incentive-compatible.

We manipulated choice architecture of the competition by randomly assigning evaluators to one of two conditions: an opt-in condition or an opt-out condition. In both conditions, we informed evaluators of the nature of the task, and told them that applicants had participated in multiple stages of a math task in a previous study, and that at one point they made a decision about whether or not to compete. In both conditions, we told the evaluators that applicants completed a math task for 5 minutes and had a tournament compensation in the first stage of the task (Stage 2 of the actual task in Study 1), whereby if the focal applicant had the highest number of correct answers among 3 other applicants, they would receive \$2 for each correct answer; if not, they received nothing. Then, in the opt-in condition, we informed the evaluators of the opt-in choice architecture into competition in the second Stage (Stage 3 of the actual task in Study 2). Conversely, in the opt-out condition, we informed evaluators of the opt-out choice frame.

Finally, evaluators in both the opt-in and opt-out conditions received information about the distribution of scores in their respective conditions. Specifically, evaluators in the opt-in and opt-out conditions in Study 2 were shown a histogram of the distribution of scores, average score, and range of scores of applicants who had been assigned to the opt-in and opt-out condition in Study 1, respectively. Following these instructions, evaluators completed two comprehension check items that asked about the nature of the task and the nature of the choice that applicants were given.

¹⁰ We chose to use Stage 2 scores as the measure of past performance, as there were slight gender differences in Stage 1 scores, and because Stage 2 was the Stage that immediately preceded Stage 3, which we used as our benchmark of a) whether the applicant entered the applicant pool and b) whether the applicant had later gone on to actually win the tournament or not.

To create the applicant pools, we selected the data of applicants who had chosen the tournament in Study 1 to be displayed in Study 2. We did this because applicants who chose the piece-rate payment in Study 1 had self-eliminated out of the applicant pool, and in a real promotion, would not be seen, much less considered, by the evaluator. We then split and matched the applicant data by condition – evaluators assigned to the opt-in condition in Study 2 would evaluate applicants who had been assigned to the opt-in condition in Study 1, and evaluators assigned to the opt-out condition in Study 2 would evaluate applicants who had been assigned to the opt-out condition in Study 1.

There were two male and two female applicants in each of the groups. The three groups differed in their composition of past performance of the two top scorers: a group where a male applicant’s past score clearly dominates a female applicant’s past score (M-high score), a group where a female applicant’s past score clearly dominates a male applicant’s past score (F-high score), and a group whereby a male applicant’s score was in a close-tie with a female applicant’s score (close-tie). We defined a close-tie as a difference of 1 or 0 between the top male and female’s scores. We presented the three groups to each evaluator in random order.

To generate these groups, our program first split the applicant dataset by condition, then took the total number of applicants who had selected the tournament in Stage 3 in each condition. The program then removed applicants who had selected the tournament but whose score was in the bottom quartile, as a way to retain only applicants who had passed a “qualification” threshold. Then, the program randomly selected two male and female applicants from the applicant pool, and classified the applicants into one of the three group types as defined above. The program continued to generate these applicant groups of two men and two women without replacement until there were no more applicants. If there were at least one group within each of the three group categories we specified (male-dominated, female-dominated, and close-tie), the program then proceeded to randomly select one of the groups within each group category to be presented to the evaluator. If there was no such group in one of the three group categories, the program took the initial applicant dataset and restarted this group creation process.

For each group, we showed the evaluators each applicant’s profiles first sequentially, then jointly in a summary page where we presented all four applicant’s profiles simultaneously. On this summary page, we asked them to indicate whom they would like to promote. For each applicant, we provided information on their gender, age, year of study, program of study, and score in Stage

2, which were taken directly from that applicant's actual demographic information that they had filled out in Study 1. Further, for each applicant, we included a brief reminder of the choice architecture of the competition. For evaluators and applicants in the opt-in condition, evaluators were told that the applicant had by default enrolled in piece-rate but had opted in to compete. For evaluators and applicants in the opt-out condition, evaluators were told that the applicant had by default enrolled in the competition and did not opt out of the competition.

After making the three promotion choices, evaluators completed a brief demographic survey that included questions about their gender, age, ethnicity, social-political orientation, economic-political orientation, and experience making hiring and promotion decisions.

3.2 Results

Our main outcome of interest is whether an evaluator chose a female applicant from each of the three 4-participant groups that they had to consider, as a function of the type of groups (male-dominated, female-dominated, or close-tie) and the choice frame condition to which the evaluators were assigned. Thus each observation in this analysis is at the evaluator-group level ($n=2,481$).

Figure 6 shows that the reported score (i.e., past performance on the math task as displayed to the evaluator) of the four applicants in a group was a strong predictor of the gender of the selected applicant. In groups where a man had a higher reported score than the best scoring woman, women were selected in about 30% of the cases; in close-tie groups, women were slightly more likely than men to be selected; finally, in groups where a woman had the top score, women were chosen about 80% of the cases. In contrast, the choice architecture condition did not relate to the gender of the chosen applicants. In the case of close-ties, the share of selected participants who were women was a little lower in the opt-out than in the opt-in condition, but the difference is not statistically significant.

In Figure 7 we focus on the close-tie groups (one observation per evaluator, $N=826$). These are the most interesting cases because they are the ones where gender bias is more likely to occur (Bohnet et al. 2016). Within the close-ties, we further distinguish three cases: the top-scoring woman in the group having a score that is lower (by one point) than the top-scoring man; the top-scoring woman having the identical score to the top-scoring man; and the top-scoring woman having a score that is higher (by one point) than the top-scoring man. The figure shows, again, that the score ranking is a strong predictor of the gender of the selected applicants, whereas the

condition (opt-in vs. opt-out) did not affect the selection. The regression estimates in Table 5 confirm the descriptive evidence. The outcome variable is a binary indicator for whether the evaluator chose a woman in a close-tie group; the main explanatory variables of interest are indicator for the different subgroups within the close-tie cases, and interactions between these indicators and an indicator for whether the particular group was assigned to the opt-out condition in Study 1. The regressions also include indicators for the socio-demographic characteristics of the evaluators that we collected, and their interactions with the opt-out dummy. The estimates confirm the strong explanatory variables of the score rank, with no difference between opt-in and opt-out choice framing. Interestingly, evaluators who self-identified as liberal on social issues were more likely, all else constant, to select a woman, and those who identified as conservative on economic issues were less likely ($p=0.1$) to select a woman in the opt-out condition. Table 6 shows that the evaluator condition did not affect the probability that the highest scorer in the stage 2 tournament was a woman.

Finally, in Table 7 we report estimates of regressions at the individual participant level, where the outcome variable is an indicator for whether a particular participant was selected from their group. The regressors include the characteristics of the participants that we included for the evaluators to review: gender, year of study, age, major, and score. We also add an indicator for whether the participant had the highest score in their group of four, an indicator for whether the participant was in the opt-out condition, and the interaction of the indicators for the female gender and the opt-out condition. The strongest predictors of the probability to be selected are being the top scorer in a group, and, conditional on being the top scorer, the raw score.

Even when controlling for individual characteristics and performance of the participants, the estimates show an evaluation bias *in favor* of promoting women in our sample. Some of the previous results corroborate this overall gender bias in favor of women; for example, in cases where there was a close or perfect tie between a man and a woman, evaluators chose the woman in more than 50% of the cases. This slant toward selecting a woman is largely due to the large proportion of evaluators who identified themselves as socially liberal. For example, if we split the sample between socially liberal on the one hand, and socially moderate or conservative on the other hand, the probability of hiring a woman is estimated to be 6.6 percentage points higher than men in the former subset of evaluators ($p<0.000$), and only 2.6 percentage points higher in the

latter subset ($p=0.073$). In neither subsample, however, does the choice architecture assigned to the evaluators have any explanatory power in the choice of a man or a woman.

This second experiment therefore shows that the choice architecture of the decision to compete did not affect the likelihood at which evaluators promote female applicants. This addressed concerns about negative unintended consequences of an opt-out promotion scheme, whereby evaluators may see women as being less motivated and desiring the promotion less. Specifically, we find that evaluators are just as likely to promote a woman under an opt-out promotion scheme compared to an opt-in promotion scheme.

4. Conclusions

Decades of research have investigated how to close the gender gap in high-level positions in organizations, to limited success. Promotion processes and the ascension to the highest ranks of an organization have the features of a competitive process, and several studies have demonstrated that women have greater aversion to competition than men. Therefore, this discrepancy in the propensity to compete has been put forth as an explanation for the gender gap in high-level organizational positions. Our study suggests that women's aversion to competition may depend on the context and, in particular, on how the option to compete is presented. Specifically, we find that an "opt-out" framing eliminates gender differences in willingness to compete.

One organizational implication that is amenable to empirical testing in the field is that changing the choice frame in the definition of applicant pools for promotions may lead to more women competing for higher positions; in turn, this would allow organizations to take advantage of a larger talent pool. In the debate between whether to "fix the women" or "fix the system", our study provides some support to the latter approach.

In addition to contributing to scholarship by providing a more nuanced view of the origins of potential aversion of women to competitive endeavors, our study also calls for further analyses to identify the mechanisms through which an opt-out frame attenuates or even eliminates gender differences in competitive choices. The choice frame may affect, for example, the perception of the type of competitive environment participants are facing. Or, different choice frames, and in particular the nature of the default option, may signal different social norms concerning competition and participation in promotion "tournaments". Individuals may also perceive a default choice to compete as more or less consistent with their gender identity, as opposed to an active

choice. Our hope is that future studies will build on the insights from our experiments to explore these questions and, by doing so, provide further insights to organizations attempting to close the gender gap in high-level positions.

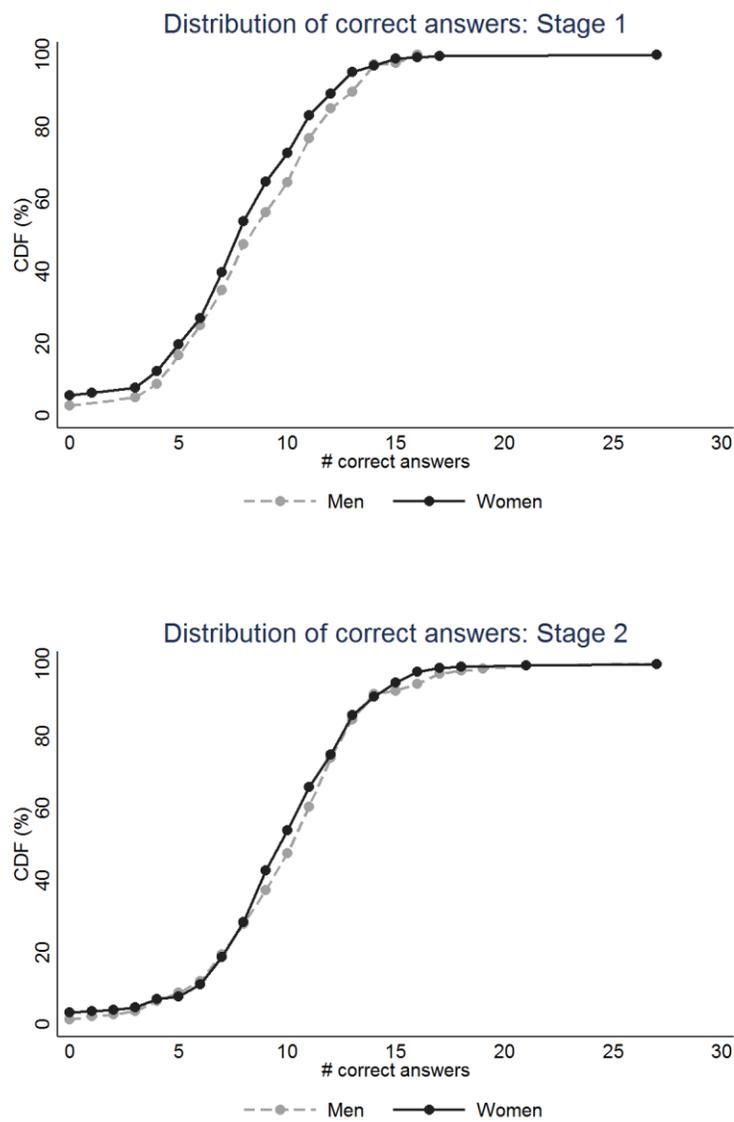
References

- Anicich EM, Fast NJ, Halevy N, Galinsky AD (2015) When the Bases of Social Hierarchy Collide: Power Without Status Drives Interpersonal Conflict. *Organ. Sci.*:orsc.2015.1019.
- BBC World Service (2018) News Hour, Donna Strickland: “Nobody expects the Nobel prize.”
- Bernardi RA, Bosco SM, Vassill KM (2006) Does Female Representation on Boards of Directors Associate With *Fortune* ’s “100 Best Companies to Work For” List? *Bus. Soc.* 45(2):235–248.
- Bohnet I (2016) *What works: Gender equality by design* (Belknap Press of Harvard University Press., Cambridge, MA).
- Bohnet I, van Geen A, Bazerman M (2016) When Performance Trumps Gender Bias: Joint vs. Separate Evaluation. *Manage. Sci.* 62(5):1225–1234.
- Bosquet C, Combes PP, García-Peñalosa C (2018) Gender and Promotions: Evidence from Academic Economists in France. *Scand. J. Econ.*
- Buhrmester M, Kwang T, Gosling SD (2011) Amazon’s Mechanical Turk. *Perspect. Psychol. Sci.* 6(1):3–5.
- Chang EH, Milkman KL, Gromet DM, Rebele RW, Massey C, Duckworth AL, Grant AM (2019) The mixed effects of online diversity training. *Proc. Natl. Acad. Sci.* 116(16):7778–7783.
- Chen DL, Schonger M, Wickens C (2016) oTree—An open-source platform for laboratory, online, and field experiments. *J. Behav. Exp. Financ.* 9:88–97.
- Choi JJ, Laibson D, Madrian BC, Metrick A (2002) Defined Contribution Pensions: Plan Rules, Participant Choices, and the Path of Least Resistance. :67–114.
- Chugh D (2018) Get from diversity to inclusion: ask 4 questions about your meetings. *Ted*. Retrieved (January 5, 2019), <https://ideas.ted.com/how-do-you-get-from-diversity-to-inclusion-ask-these-4-questions-about-your-meetings/>.
- Daubman KA, Heatherington L, Ahn A (1992) Gender and the self-presentation of academic achievement. *Sex Roles* 27(3–4):187–204.
- DeCelles KA, DeVoe SE, Rafaeli A, Agasi S (2019) Helping to reduce fights before flights: How environmental stressors in organizations shape customer emotions and customer-employee interactions. *Pers. Psychol.* 72(1):49–80.
- Dezsö CL, Ross DG (2012) Does female representation in top management improve firm performance? A panel data investigation. *Strateg. Manag. J.* 33(9):1072–1089.
- Dobbin F, Kalev A (2016) Why Diversity Programs Fail. *Harv. Bus. Rev.* 94(7).
- Downs JS, Loewenstein G, Wisdom J (2009) Strategies for Promoting Healthier Food Choices. *Am. Econ. Rev.* 99(2):159–164.

- Exley CL, Niederle M, Vesterlund L (2016) Knowing When to Ask: The Cost of Leaning In. *Exp. Econ.* 10(2):171–178.
- Fischbacher U (2007) z-Tree: Zurich toolbox for ready-made economic experiments. *Exp. Econ.* 10(2):171–178.
- Fitzsimons GM, Kay AC, Kim JY (2018) “Lean In” Messages and the Illusion of Control. *Harv. Bus. Rev.*
- Flory JA, Leibbrandt A, List JA (2015) Do Competitive Workplaces Deter Female Workers? A Large-Scale Natural Field Experiment on Job Entry Decisions. *Rev. Econ. Stud.* 82(1):122–155.
- Harmatz MG, Lapuc PS (1968) A technique for employing a yoked control in free operant verbal conditioning experiments. *Behav. Res. Ther.* 6(4):483.
- Healy A, Pate J (2011) Can Teams Help to Close the Gender Competition Gap? *Econ. J.* 121(555):1192–1204.
- Johnson EJ, Bellman S, Lohse GL (2002) Defaults, Framing and Privacy: Why Opting In-Opting Out. *Mark. Lett.* 13(1):5–15.
- Johnson EJ, Goldstein D (2003) Do defaults save lives? *Science* (80-.). 302(5649):1338–1339.
- Johnson SK, Murphy SE, Zewdie S, Reichard RJ (2008) The strong, sensitive type: Effects of gender stereotypes and leadership prototypes on the evaluation of male and female leaders. *Organ. Behav. Hum. Decis. Process.* 106(1):39–60.
- Kalev A, Dobbin F, Kelly E (2006) Best Practices or Best Guesses? Assessing the Efficacy of Corporate Affirmative Action and Diversity Policies. *Am. Sociol. Rev.* 71(4):589–617.
- Kim JY, Fitzsimons GM, Kay AC (2018) Lean in messages increase attributions of women’s responsibility for gender inequality. *J. Pers. Soc. Psychol.* 115(6):974–1001.
- Lazear EP, Rosen S (1981) Rank-Order Tournaments as Optimum Labor Contracts. *J. Polit. Econ.* 89(5):841–864.
- Leslie LM (2018) Diversity Initiative Effectiveness: A Typological Theory of Unintended Consequences. *Acad. Manag. Rev.*:amr.2017.0087.
- Macrae CN, Bodenhausen G V., Milne AB, Jetten J (1994) Out of mind but back in sight: Stereotypes on the rebound. *J. Pers. Soc. Psychol.* 67(5):808–817.
- Marteau TM, Bekker H (1992) The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *Br. J. Clin. Psychol.* 31(3):301–306.
- Mols F, Haslam SA, Jetten J, Steffens NK (2015) Why a nudge is not enough: A social identity critique of governance by stealth. *Eur. J. Polit. Res.* 54(1):81–98.
- Moss-Racusin CA, Rudman LA (2010) Disruptions in Women’s Self-Promotion: The Backlash Avoidance Model. *Psychol. Women Q.* 34(2):186–202.
- Niederle M, Segal C, Vesterlund L (2013) How Costly Is Diversity? Affirmative Action in Light of Gender Differences in Competitiveness. *Manage. Sci.* 59(1):1–16.
- Niederle M, Vesterlund L (2007) Do Women Shy Away From Competition? Do Men Compete Too Much? *Q. J. Econ.* 122(3):1067–1101.
- Reuben E, Rey-Biel P, Sapienza P, Zingales L (2012) The emergence of male leadership in competitive environments. *J. Econ. Behav. Organ.* 83(1):111–117.
- Rudman LA (1998) Self-promotion as a risk factor for women: The costs and benefits of counterstereotypical impression management. *J. Pers. Soc. Psychol.* 74(3):629–645.
- Rudman LA, Glick P (2001) Prescriptive Gender Stereotypes and Backlash Toward Agentic Women. *J. Soc. Issues* 57(4):743–762.
- Samek A (2017) Management Science Gender Differences in Job Entry Decisions: A University-Wide Field Experiment Gender Differences in Job Entry Decisions: A University-Wide

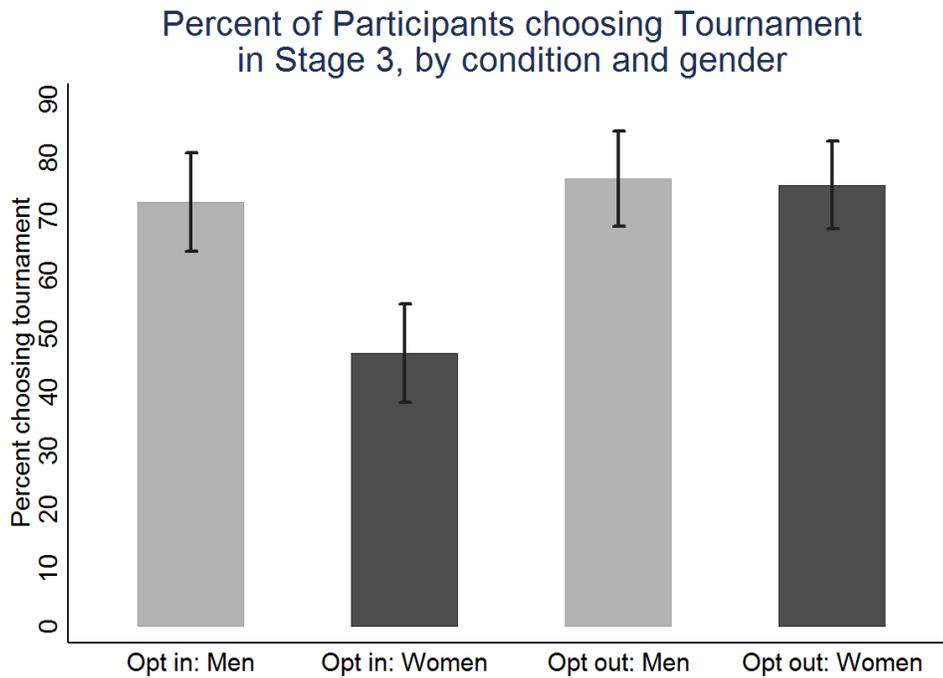
- Field Experiment. :1–10.
- Samuelson W, Zeckhauser R (1988) *Status Quo Bias in Decision Making*
- Sanchez JI, Medkik N (2004) The Effects of Diversity Awareness Training on Differential Treatment. *Gr. Organ. Manag.* 29(4):517–536.
- Sandberg S (2013) *Lean in: Women, work, and the will to lead.* (Knopf Doubleday Publishing Group).
- Selinger E, Whyte K (2011) Is There a Right Way to Nudge? The Practice and Ethics of Choice Architecture. *Sociol. Compass* 5(10):923–935.
- Spielberger C, Gorsuch R (1983) *State-trait anxiety inventory for adults: Manual and sample: Manual, instrument and scoring guide* (Consulting Psychologists Press).
- Thaler RH, Sunstein CR (2008) *Nudge: improving decisions about health, wealth, and happiness* (Yale University Press).
- Tluczek A, Henriques JB, Brown RL (2009) Support for the reliability and validity of a six-item state anxiety scale derived from the State-Trait Anxiety Inventory. *J. Nurs. Meas.* 17(1):19–28.
- Weber A, Zulehner C (2010) Female Hires and the Success of Start-up Firms. *Am. Econ. Rev.* 100(2):358–361.

Figure 1: In Study 1, the distribution of correct answers in Stages 1 and 2 did not differ between men and women



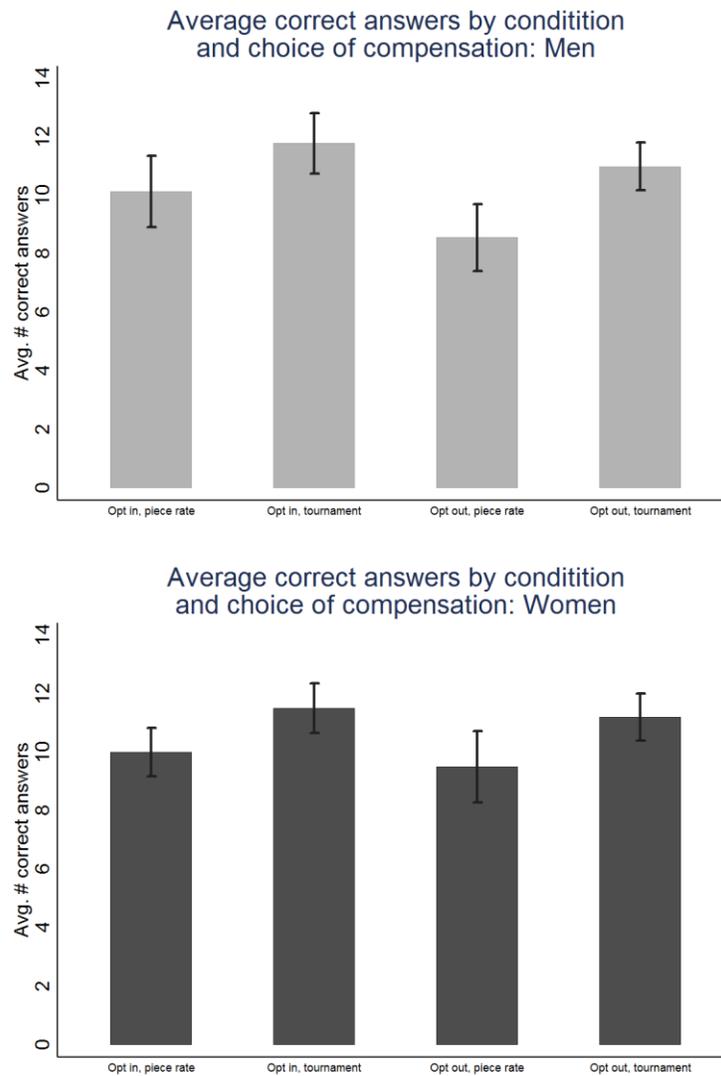
Notes: The graphs report the observed cumulative distribution of correct responses by men and women in stages 1 and 2 of the experiment.

Figure 2: In Study 1, the opt-out choice framework for the participation in a tournament in Stage 3 eliminated the gender difference in the choice to compete in the tournament



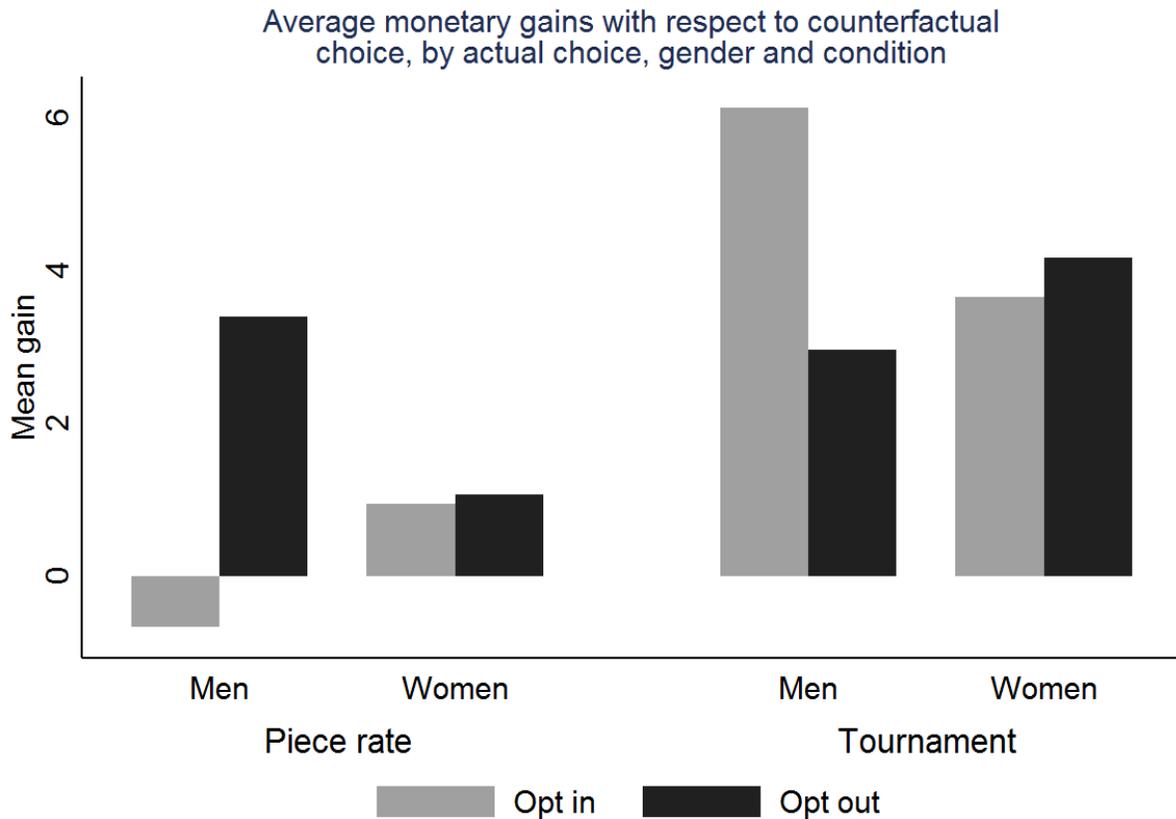
Notes: “Opt in” indicates that participants were in the condition where the default compensation was piece rate. “Opt out” indicates the condition where participants were assigned to competing in a winner-take all tournament within each group of four by default, but could opt-out and instead be compensated on a piece rate basis. The vertical lines indicate 95% confidence interval for the mean percentages.

Figure 3: In Study 1, average correct responses in Stage 3 did not differ by condition, within a chosen compensation method



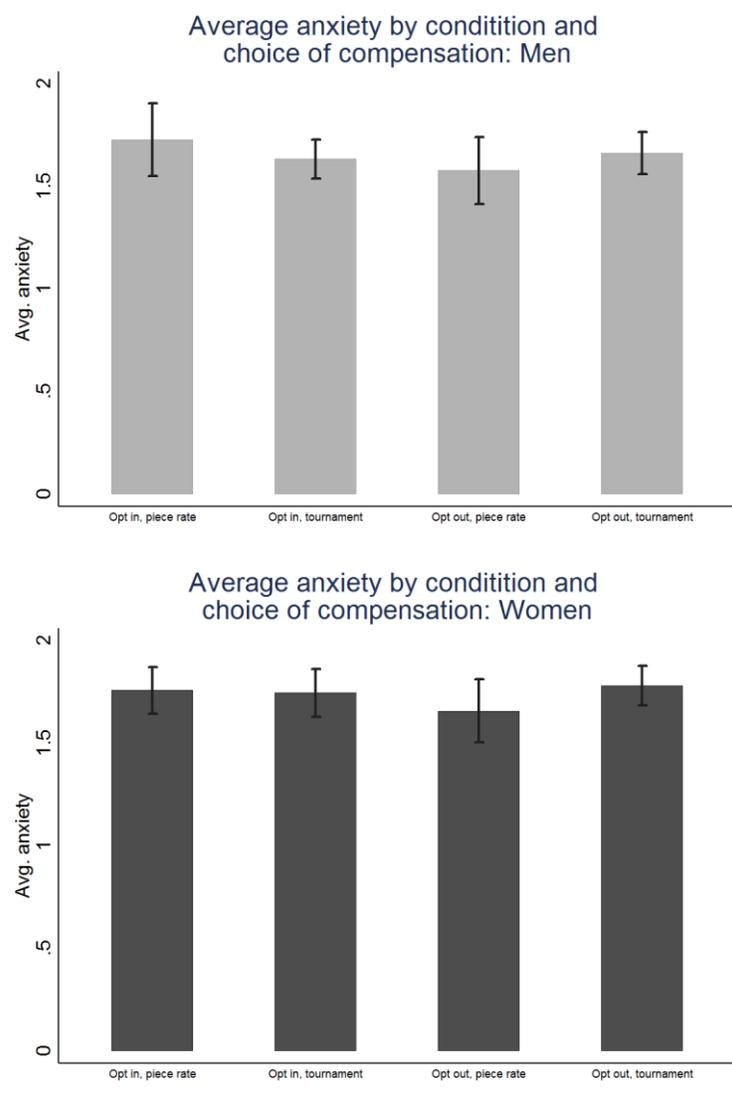
Notes: The vertical lines indicate 95% confidence interval for the means.

Figure 4: In Study 1, women in the opt-out condition achieved higher average monetary gains than men did; the overall additional monetary gains were more equally distributed between men and women in the opt-out compared to the opt-in condition.



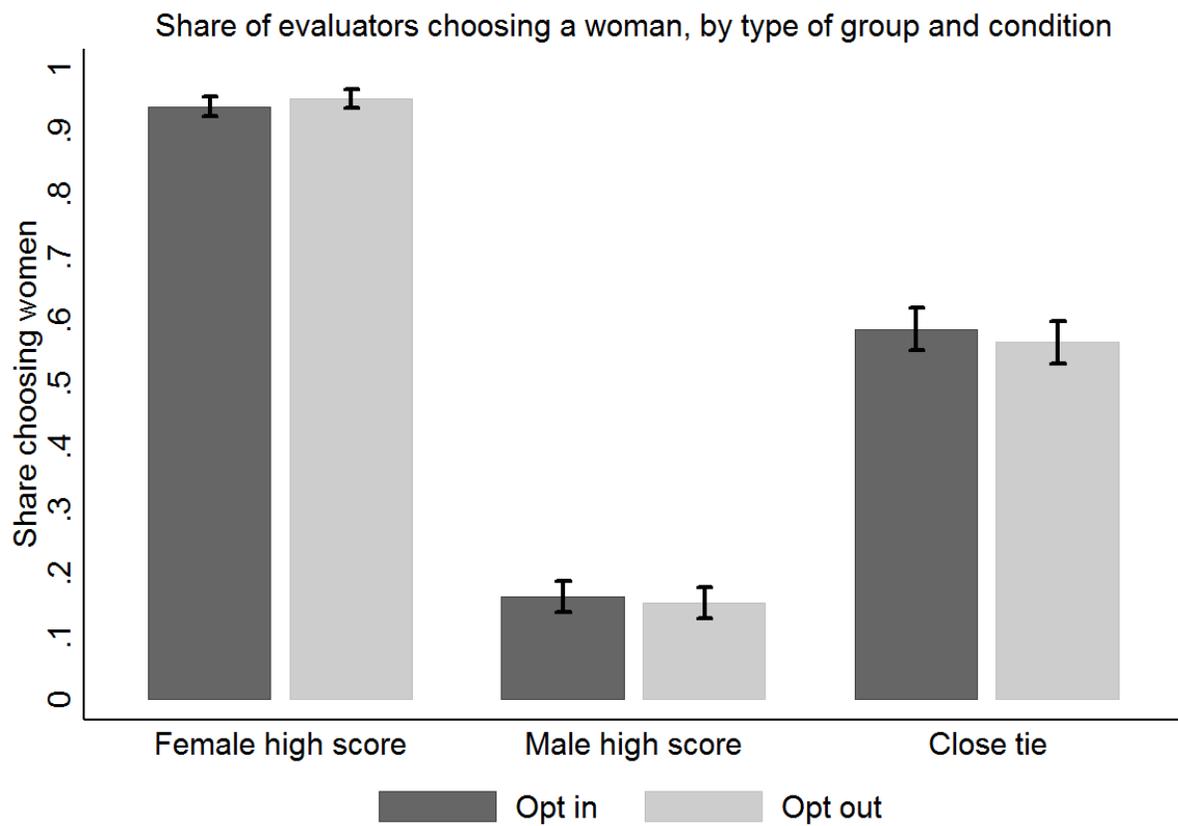
Notes: The graph reports the average difference between the average dollar payoff that participants realized in Stage 3 and the payoff that they would have realized if they chose the alternative compensation scheme, given the number of correct answers they gave in stage 3. Each column represents a combination of gender, the actual choice of the participant, and the experimental condition (Opt in vs. Opt out). For the participants who selected a tournament-based compensation scheme, the “counterfactual” payoff is the number of correct responses that they gave in Stage 3 multiplied by \$0.5. For the participants who selected piece rate, the counterfactual payoff is zero if they would have not won the tournament in the group to which they would be assigned, and equal to the number of their correct answers multiplied by \$2 had they been the winners of the groups to which they were assigned.

Figure 5: In Study 1, average anxiety levels were slightly higher for women, but did not differ by condition and choice of compensation



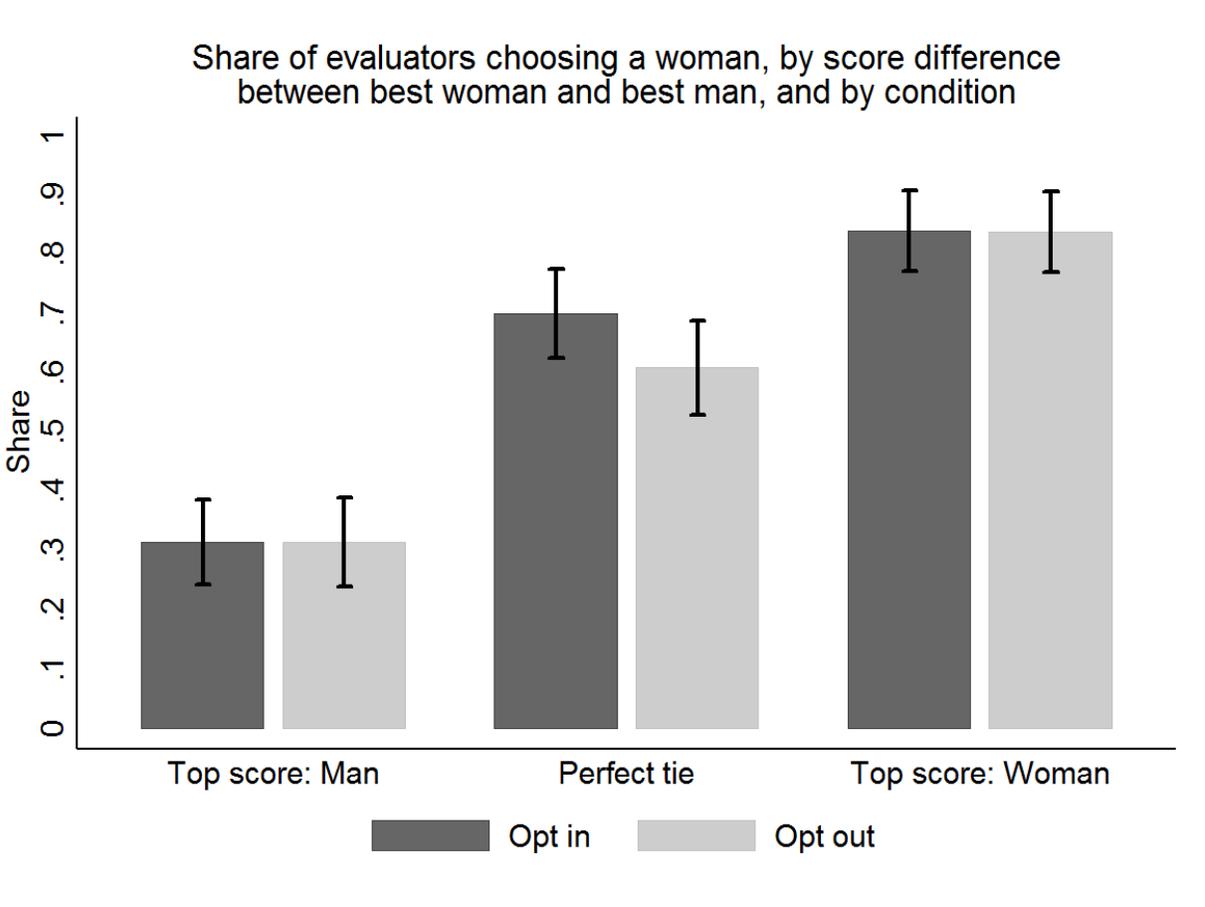
Notes: The measure of anxiety is the average of the scores (integer numbers between 1 and 4) that each respondent assigned to five questions of the State Anxiety Inventory (SAI) (Marteau and Bekker 1992), adapted from the original scale (Spielberger and Gorsuch 1983). A higher value indicates a higher anxiety score.

Figure 6: In Study 2, the reported score of the four contestants in a group were a strong predictor of the gender of the selected participants



Notes: “Female (Male) high score dominated” groups indicate groups of four contestants where a woman (man) was the contestant with the highest reported score, and this score was strictly greater than the score of the best-performing man (woman) in that group. “Close-tie” groups are the groups of four contestant in which a woman and a man were both the highest performers (equal reported score), or a woman (man) had the highest reported score, and a man (woman) had a reported score that was 1 point lower.

Figure 7: In Study 2, even within “close-tie” groups, score was strong predictor of the gender of the selected applicants, whereas the condition (opt-in vs. opt-out) did not affect the selection



Notes: “Top score: (Man)” indicated the close-tie groups where a man (woman) had the highest reported score, and a woman (man) had a reported score that was 1 point lower. “Perfect tie” indicated the groups where two contestants, a man and a woman, had the same, highest reported score. The vertical lines indicate 95% confidence interval for the means.

Table 1: Study 1 - Regression analyses of the choice of tournament-based compensation in Stage 3

Outcome variable:	Choice of tournament in stage 3				
	(1)	(2)	(3)	(4)	(5)
Woman	-0.137*** (0.045)				
Opt in: Woman		-0.254*** (0.070)	-0.241*** (0.071)	-0.247*** (0.075)	-0.246*** (0.072)
Opt out: Man		0.043 (0.068)	0.053 (0.068)	0.030 (0.072)	0.040 (0.069)
Opt out: Woman		0.032 (0.072)	0.038 (0.074)	0.033 (0.076)	0.035 (0.073)
# correct answ. in stage 2			0.021*** (0.007)	0.004 (0.009)	0.025*** (0.007)
# correct in stage 2 - # correct in stage 1			-0.004 (0.009)	-0.009 (0.008)	-0.005 (0.009)
Guessed rank in stage 2 tournament				-0.132*** (0.033)	
(Over)confidence					0.044** (0.022)
Observations	482	482	482	482	482
Pseudo R2	0.0168	0.0558	0.0745	0.115	0.0823

Notes: The estimates indicate marginal effects from probit regressions. The outcome variable is a binary indicator equal to 1 if a participant selected tournament based compensation in stage 3, and 0 if they selected piece rate compensation. Regressors include: the experimental conditions (opt-in vs. opt-out frame) interacted with the gender of the participant (the omitted category is men in the opt-in condition); the number of correct responses in stage 2, and the difference between the correct answers in stage 1 and those in Stage 1; the positions that each participant guessed to have achieved in the tournament in stage 2 (out of four position, rank 1 being the winner); and the difference between the actual position and the guessed position, as a measure of (over) confidence. The baseline is a male participant in the opt-in condition, with 8.52 correct answers in stage 1, 10.24 correct answers in stage 2, a guess for their rank in stage 2 of 2.05, and a confidence (actual rank *minus* guessed rank) of 0.31. The share of mean in the opt-in conditions who chose a tournament is 0.725 (or 72.5%). Estimated standard errors, clustered at the session level (there were 36 sessions) are in parentheses. * p<0.1, ** p<0.05, *** p<0.01 (two-sided tests).

Table 2: Study 1 - Number and share of payoff-maximizing choices by gender, choice and condition

	Number of payoff-maximizing choices	Percent of payoff maximizing choices	N
Opt in, piece rate: Men	23	76.7%	30
Opt in, tournament: Men	38	48.1%	79
Opt out, piece rate: Men	23	92.0%	25
Opt out, tournament: Men	34	42.0%	81
Opt in, piece rate: Women	58	79.5%	73
Opt in, tournament: Women	23	35.9%	64
Opt out, piece rate: Women	25	78.1%	32
Opt out, tournament: Women	48	49.0%	98

Notes: The table reports the number and percentage, of subjects who made the payoff maximizing choice for them in Stage 3 by experimental condition, gender, and choice of compensation scheme. We established the payoff maximizing choice as follows. We used the estimated coefficients from a probit regression of whether a participant won their tournament in Stage 2 on the number of correct responses in that stage, to predict the likelihood of winning a tournament in Stage 3 given their performance in Stage 3 (recall that individuals who chose the tournament in Stage 3 had their performance compared against three participants from Stage 2). We then calculated the *expected* payoff from choosing a tournament or a piece rate compensation for each participant. For the piece rate, the expected payoff was \$0.50 X the number of correct responses in Stage 3. The expected payoff from a tournament was \$2 X the number of correct responses in Stage 3 X the predicted probability of winning the tournament for that subject in Stage 3. We classify a participant as having made their payoff-maximizing choice if they selected the compensation scheme that gave them the higher expected payoff. In cases where the expected payoffs from the two compensation scheme were close to each other (less than \$1 in absolute difference; this happened if a participant solved 12 questions correctly), we randomized the assignment to having made the payoff maximizing decision or not (we interpreted small differences as making a participant indifferent between the two schemes).

Table 3: Study 1 - Average monetary gains (losses) from choosing a compensation scheme in Stage 3, compared to “counterfactual” choice

	Sum of net gains	Avg. net gains	N
<i>Gender</i>			
Men	\$790.5	\$3.68	215
Women	\$746.0	\$2.79	267
<i>Choice architecture</i>			
Opt in	\$767.5	\$3.12	246
Opt out	\$769.0	\$3.26	236
<i>Conditions in stage 3, by gender</i>			
Opt in: Men	\$465.0	\$4.27	109
Opt in: Women	\$302.5	\$2.21	137
Opt out: Men	\$325.5	\$3.07	106
Opt out: Women	\$443.5	\$3.41	130
<i>Conditions and compensation choices in stage 3, by gender</i>			
Opt in, piece rate: Men	-\$20.0	-\$0.67	30
Opt out, piece rate: Men	\$85.0	\$3.40	25
Opt in, piece rate: Women	\$68.5	\$0.94	73
Opt out, piece rate: Women	\$34.0	\$1.06	32
Opt in, tournament: Men	\$485.0	\$6.14	79
Opt out, tournament: Men	\$240.5	\$2.97	81
Opt in, tournament: Women	\$234.0	\$3.66	64
Opt out, tournament: Women	\$409.5	\$4.18	98

Notes: this table reports both the total gains (over all participants) and the average gains per participants from choosing a compensation scheme over the alternative one. For the participants who selected a tournament-based compensation scheme, the “counterfactual” payoff is the number of correct responses that they gave in Stage 3 multiplied by \$0.5. For the participants who selected piece rate, the counterfactual payoff is zero if they would have not won the tournament in the group to which they would be assigned, and equal to the number of their correct answers multiplied by \$2 had they been the winners of the groups to which they were assigned.

Table 4: Study 1 - Additional specification for compensation choice regressions.

Outcome variable:	Choice of tournament in stage 3
Opt in: Woman	-0.444 (0.308)
Opt out: Man	-0.006 (0.375)
Opt out: Woman	-0.491 (0.329)
# correct answ. in stage 2	0.001 (0.021)
# correct in stage 2 - # correct in stage 1	-0.013 (0.020)
Guessed rank in stage 2 tournament	-0.199*** (0.061)
# correct answ. in stage 2	0.006 (0.023)
x Opt in: Woman	-0.003 (0.028)
# correct answ. in stage 2	0.006 (0.025)
x Opt out: Woman	-0.006 (0.030)
# correct in stage 2 - # correct in stage 1	-0.005 (0.033)
x Opt in: Woman	0.037 (0.027)
# correct in stage 2 - # correct in stage 1	0.063 (0.064)
x Opt out: Man	0.030 (0.076)
Guessed rank in stage 2 tournament	0.176** (0.086)
x Opt out: Woman	
Observations	482
Pseudo R2	0.129

Notes: The estimates indicate marginal effects from probit regressions. The outcome variable is a binary indicator equal to 1 if a participant selected tournament based compensation in stage 3, and 0 if they selected piece rate compensation. Regressors include the experimental conditions (opt-in vs. opt-out frame) interacted with the gender of the participant (the omitted category is men in the opt-in condition); the number of correct responses in stage 2, and the difference between the correct answers in stage 2 and those in stage 1; the positions that each participant guessed to have achieved in the tournament in stage 2 (out of four position, rank 1 being the winner); and interaction of the gender of the participant with correct responses in Stage 2, and the difference between the correct answers in stage 1 and those in Stage 1; the positions that each participant guessed to have achieved in the tournament in stage 2. The baseline is a male participant in the opt-in condition, with 8.52 correct answers in Stage 1, 10.24 correct answers in Stage 2, a guess for their rank in stage 2 of 2.04, and a confidence (actual rank *minus* guessed rank) of 0.31. Estimated standard errors, clustered at the session level (there were 36 sessions) are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 5: Study 2 Regression analyses of the determinants of the selection of a female participant by evaluators

Outcome variable: Sample: Regression specification:	Choice of a female participant			
	Only close-tie groups		OLS	Probit
	OLS (1)	Probit (2)	OLS (3)	Probit (4)
Regressors:				
<i>Features of study 1 and Study 1 participants</i>				
Opt-out condition	-0.092* (0.056)	-0.097* (0.059)	-0.087 (0.055)	-0.096 (0.059)
Top score: Man	-0.386*** (0.053)	-0.384*** (0.053)	-0.380*** (0.053)	-0.388*** (0.054)
Top score: Woman	0.139*** (0.052)	0.175*** (0.064)	0.142*** (0.051)	0.180*** (0.064)
Opt-out condition X Top score: Man	0.092 (0.077)	0.095 (0.078)	0.095 (0.076)	0.103 (0.079)
Opt-out condition X Top score: Woman	0.090 (0.074)	0.092 (0.092)	0.086 (0.074)	0.094 (0.093)
Controls (characteristics of evaluators)			X	X
Constant	0.701*** (0.038)		0.748*** (0.054)	
Observations	826	826	826	826
R-squared	0.195		0.218	
Pseudo R2		0.150		0.171

Notes: The outcome variable is a binary indicator equal to 1 if an evaluator selected a woman from a group of four participants. The sample includes only groups of participants with a “close-tie”, i.e. where the difference in score (in Study 1) between the highest performing man and the highest performing woman was -1, 0 and 1. The variable “Opt-out condition” is equal to 1 if the participants in a group were, in study 1, assigned to the Opt-out condition, and 0 otherwise. “Top score: Man (Woman)” is equal to 1 if, in a given group, the top scoring man (woman) had one more correct answer than the top scoring woman (man), and zero otherwise; the omitted category is, therefore, the groups with a perfect tie between the top scoring man and woman. Control variables include the various characteristics of the evaluators that we collected: gender, age, educational attainment, social and economic orientations, and previous experience in HR work. For the Probit specifications, we report the estimates of the marginal effects. Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 6: Study 2 - Regression analyses of the determinants of the selection of the top scoring female participant by evaluators, conditional on selecting a woman

Outcome variable:	Choice of the woman with the highest score in Stage 2, conditional on choosing a woman				
	Sample:	Full sample		Close-tie groups	
		Regression specification:	OLS (1)	Probit (2)	OLS (3)
Regressors					
Opt-out condition		-0.004 (0.045)	-0.005 (0.043)	-0.068 (0.073)	-0.068 (0.074)
Male-dominated group		-0.339*** (0.055)	-0.381*** (0.069)		
Female-dominated group		0.280*** (0.038)	0.295*** (0.039)		
Opt-out condition X Male-dominated group		-0.015 (0.079)	-0.024 (0.114)		
Opt-out condition X Female-dominated group		-0.022 (0.054)	-0.028 (0.058)		
Top score: Man				-0.006 (0.087)	-0.004 (0.088)
Top score: Woman				0.036 (0.073)	0.037 (0.074)
Opt-out condition X Top score: Man				-0.013 (0.124)	-0.021 (0.128)
Opt-out condition X Top score: Woman				0.155 (0.102)	0.162 (0.100)
Constant		0.536*** (0.045)		0.567*** (0.076)	
Observations		1,390	1,390	477	477
R-squared		0.178		0.065	
Pseudo R2			0.139		0.049

Notes: The outcome variable is a binary indicator equal to 1 if an evaluator selected a woman and that woman was the top scorer (winner) of the Stage 3 tournament in Study 1. The variable “Opt-out condition” is equal to 1 if the participants in a group were, in study 1, assigned to the Opt-out condition, and 0 otherwise. “Male (Female) – dominated group” is equal to 1 if in a given group, a man (woman) had a score from Stage 2 of Study 1 that was greater than the one of the best scoring woman (man) by 2 or more. The omitted category includes the close-tie groups. “Top score: Man (Woman)” is equal to 1 if, in a given close-tie group, the top scoring man (woman) had one more correct answer than the top scoring woman (man), and zero otherwise; the omitted category is, therefore, the groups with a perfect tie between the top scoring man and woman. Control variables include the various characteristics of the evaluators that we collected: gender, age, educational attainment, social and economic orientations, and previous experience in HR work. For the Probit specifications, we report the estimates of the marginal effects. Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 7: Study 2 - Regression analyses of the determinants of the selection of a participant

Outcome variable: Sample: Regression specification:	Participant is selected							
	Full sample				Socially moderate and conservative evaluators		Socially liberal evaluators	
	OLS	Probit	OLS	Probit	OLS	Probit	OLS	Probit
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Regressors :</i>								
Woman	0.064*** (0.012)	0.064*** (0.012)	0.048*** (0.008)	0.067*** (0.013)	0.026* (0.015)	0.029 (0.022)	0.066*** (0.012)	0.102*** (0.021)
Opt-out condition	0.003 (0.012)	0.003 (0.013)	0.009 (0.008)	0.013 (0.013)	0.009 (0.011)	0.010 (0.017)	0.009 (0.010)	0.015 (0.017)
Woman X Opt-out condition	-0.006 (0.017)	-0.006 (0.017)	-0.012 (0.012)	-0.013 (0.018)	-0.009 (0.021)	-0.005 (0.031)	-0.013 (0.017)	-0.017 (0.027)
Score in Stage 1			0.017*** (0.001)	0.020*** (0.002)	0.015*** (0.002)	0.016*** (0.003)	0.019*** (0.002)	0.025*** (0.003)
Had highest score in Stage 1			0.626*** (0.008)	0.620*** (0.012)	0.599*** (0.020)	0.595*** (0.021)	0.647*** (0.017)	0.641*** (0.018)
Controls (characteristics of the participants)			X	X	X	X	X	X
Constant	0.218*** (0.009)		-0.446 (0.276)		-0.592 (0.500)		-0.250 (0.384)	
Observations	9,912	9,912	9,912	9,869	4,584	4,558	5,256	5,239
R-squared	0.005		0.552		0.501		0.600	
Pseudo R2		0.004		0.490		0.438		0.544

Notes: The outcome variable is a binary indicator equal to 1 if a participant was selected by an evaluator. “Woman” has value of 1 if a participant was a woman, and zero otherwise. “Opt-out condition” is equal to 1 if the participant was, in Study 1, assigned to the Opt-out condition, and 0 otherwise. “Score in stage 1” is equal to the number of correct responses that a participant had in the first tournament competition in Study 1, as reported to the evaluators. “Had highest score in Stage 1” is equal to 1 if the participant was the top scorer among the four participants in a given group. Control variables include the various characteristics of the participants that we reported to the evaluators: age, year of study, and major. For the Probit specifications, we report the estimates of the marginal effects. Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.