The impact of role models on women's self-selection in competitive environments

Kristina Meier, Alexandra Niessen-Ruenzi, and Stefan Ruenzi*

December 2018

Abstract

We show that female role models increase women's willingness to compete. As in Niederle and Vesterlund (2007), we find that women are less willing to enter a tournament than men, although there are no gender differences in performance. However, the gender gap in tournament entry disappears if subjects are exposed to a competitive female role model. Results are stronger for the best performing women who seem to be particularly encouraged by female role models. Female role models also mitigate gender stereotype threats and lead to higher self-confidence among women. By contrast, we find that competitive male role models seem to intimidate female subjects and increase the gender gap in tournament entry even further. Our results have implications for the socio-political debate on how the fraction of women in top management positions can be increased.

JEL-Classification Codes: J16, C9

Keywords: Gender, Competition, Role Models

---

*Kristina Meier can be reached at kmeier@uni-mannheim.de, Alexandra Niessen-Ruenzi can be reached at niessen@uni-mannheim.de, and Stefan Ruenzi can be reached at ruenzi@uni-mannheim.de. All authors are at the University of Mannheim. Please address all correspondence to: Alexandra Niessen-Ruenzi, University of Mannheim, L9, 1-2, 68131 Mannheim, Germany, tel. +49-(0)621-181-1595. We would like to thank the Dr. Werner Jackstädt Foundation for financially supporting this research study. We also thank Muriel Niederle, seminar participants at the Center of European Studies (ZEW), the University of Mannheim, and the University of Glasgow as well as participants at the 11th M-BEES conference, the SABE-IAREP conference, and the EFA and EEA meetings for helpful comments. All errors are our own.

1 Introduction

Although several countries have introduced gender quotas to promote a higher representation of women in top management positions, the fraction of female top managers is still very low. For example, based on a sample of board data for listed firms in 20 countries from 2001-2010, Adams and Kirchmaier (2016) show that the fraction of women on boards ranges between 7 percent and 9 percent globally and between 9 percent and 11 percent in the United States.

One suggested reason why women are under-represented in top management positions is that they are reluctant to compete against others. According to Gneezy, Niederle, and Rustichini (2003) and Niederle and Vesterlund (2007), there are two possible causes for a gender gap in the willingness to compete: First, women may be subject to a stereotype threat. If there is a stereotype according to which their gender is expected to underperform in competitions, the fear of confirming this stereotype eventually leads to lower self-efficacy and a lower willingness of women to compete (Steele and Aronson (1995)).[1] Second, being competitive may not be a socially desirable trait for women. For example, Bursztyn, Egorov, and Fiorin (2017) find that single female MBA students will downplay their career ambitions in front of male classmates presumably because they assume that men prefer less ambitious women (Günther, Ekinci, Schwieren, and Strobel (2010), Fisman, Iyengar, Kamenica, and Simonson (2006)). Similarly, Cadsby, Servátka, and Song (2013) show that women exhibit a much greater preference for competition when primed with a professional identity rather than with a gender/family identity.

[1] Evidence for this phenomenon has been brought forward with respect to women and math performance (Spencer et al. 1999), African Americans and tests of intellectual capabilities (Steele und Aronson 1995; Marx und Goff 2005), as well as European-Americans and athletic ability (Stone et al. 1999).

Since top management positions are usually characterized by tough competition among managers, these positions might thus be less attractive for women. For example, Flory, Leibbrandt, and List (2015) find that women are less likely to apply for jobs with a competitive compensation scheme, eventually self-selecting into other segments of the labor market. A number of experimental studies confirm this result by showing that women self-select into less competitive environments, especially when having to compete against men and that their performance suffers in competitive environments compared to that of men. These results are obtained despite the fact that there were no performance difference between female and male subjects in the employed tasks. For a review of this literature, refer to Niederle and Vesterlund (2011) and Niederle (2015).

In this paper, we investigate whether competitive female role models increase women's willingness to compete. The preference against competition seems to be at least partially due to culture and norms (Gneezy, Leonard, and List (2009)). Past research has shown that role models can have a short and long term impact on both stereotype threat and social values. For example, several studies find that in-group role models can increase performance in situations where stereotype threat is present (Marx and Roman (2002); Marx, Monroe, Cole, and Gilbert (2013); Marx, Ko, and Friedman (2009); Stout, Dasgupta, Hunsinger, and McManus (2011)).[2] Based on this literature, we argue that women might be more willing to participate in tournaments if they have observed another woman successfully participating in tournaments and talking positively about their experience with competition.

We adapt the experimental design of Niederle and Vesterlund (2007), which consists of solving as many basic math problems in a short period of time as possible.

[2]In addition, while social norms are typically thought to be long-lasting, recent evidence suggests that, for example, prominent political figures can also change values and behaviors (Bursztyn, Fujiwara, and Pallais (2017); Beaman, Duflo, Pande, and Topalova (2012)).

Although gender differences in mathematics performance are insignificant for most countries, stereotypes about female inferiority in mathematics may lead to female underperformance and anxiety via stereotype threat (Else-Quest, Hyde, and Linn (2010)). The gender stereotype threat in this setting may come from two different sources simultaneously: The first one is derived from the view that women are not as good in math as men, and the second one is derived from the view that women are not as competitive as men (Niederle and Vesterlund (2007)).

In our experiment, we closely follow the experimental setting suggested in Niederle and Vesterlund (2007) and ask subjects to add up five two-digit numbers, a task for which we can show that no gender differences in ability exist among our participants. The first round of the experiment involves a piece-rate compensation of 50 Cents for every correct answer a subject has given. The second round of the experiment involves a tournament compensation. Subjects are pooled in groups with four members (two female and two male members, respectively) and only the best member is paid $4 for each correct answer, while the other group members receive no payment. Before the beginning of the third round, subjects can choose whether they want to be compensated by the piece-rate or the tournament scheme for their performance in this round.

The experiment is conducted on Amazon's Mechanical Turk (AMT) platform and comprises 847 observations. In contrast to Niederle and Vesterlund (2007), we allow subjects to use calculators, because they cannot be monitored. Thus, we are in fact using a real-effort task which is only framed as a math tournament, while no actual math skills are needed to solve the calculation problems. Still, we expect gender stereotypes that have been documented in the literature (Bhana (2005), Fennema, Peterson, Carpenter, and Lubinski (1990)) to be strong enough to lead to gender differences in tournament entry.

To test our hypothesis that female role models mitigate the gender gap in tournament entry, we use a 3 (female role model, male role model, no role model) x 2 (subject gender) between-subject design. Our experiment comprises three conditions and subjects are randomly sorted into one of them. In the "neutral" condition, no role models are shown. We use this condition to replicate the results of Niederle and Vesterlund (2007) and to make sure that our sample delivers the same baseline result. Our results in this condition show that women indeed shy away from competition and thus confirm their results.

In the female role model condition, we show a female role model video, which subjects watch before they decide whether they want to be compensated according to the tournament scheme or the piece-rate scheme. The videos consist of an interview with a successful woman who talks positively about competition and how much she enjoys to compete. In this condition, more female subjects chose the tournament, and the gender difference in tournament entry is not statistically significant anymore.

By contrast, in the male role model condition, where we show a male role model video, gender difference get even more pronounced as compared to the neutral condition. Thus, male role models seem to strongly activate gender stereotypes and intimidate female subjects, who shy away from competition even more than if they observe no role model at all.

Interestingly, we find that performance in the first two rounds predicts tournament entry for male subjects, but not for female subjects. That is, better performing men are more likely to enter the tournament, while we find no such effect for women. However, in line with the finding in Carrell, Page, and West (2010) that female professors have the largest impact on the performance and college majors of female students with high prior math ability, the impact of female role models on female subjects' tournament entry decision is strongest for the best performing women. These women are significantly more likely to enter the tournament in the female role model condition compared to the male role model and neutral conditions. As in Carrell, Page, and West (2010), we do not find a significant impact of female role models on female subjects in the worst performing quartile of the distribution. This finding suggests that the best performing women seem to be hindered most from tournament entry by gender stereotypes and benefit most from encouragement through female role models. Thus, female role models lead to a higher fraction of those women in the tournament condition, who are most likely to win it, and thus should enter it in the first place.

This paper contributes to several strands of the literature. First, we contribute to the literature on the impact of female role models on women's self selection into competitive environments (Bettinger and Long (2005), Stout, Dasgupta, Hunsinger, and McManus (2011), Beaman, Duflo, Pande, and Topalova (2012)). Bettinger and Long (2005) investigate the potential impact of female faculty as role models on students' choice of subjects. They show that female students chose mathematical subjects more frequently if they already took a math class with a female instructor before. Similarly, Stout, Dasgupta, Hunsinger, and McManus (2011) show that exposure to female experts in Science, Technology, Engineering, and Mathematics (STEM) leads to more positive attitudes and self-efficacy in STEM among women. However, in these settings, women have already self selected into math classes or STEM when their behavior and attitudes are studied. In contrast, our study investigates the impact of female role models on women *before* they self-select in a competitive environment. Based on field and survey data, Beaman, Duflo, Pande, and Topalova (2012) show that gender differences in career aspirations and educational attainment decrease if adolescents live in Indian villages where a female leader is appointed, which also points towards a role model effect. Our paper augments these results by providing direct experimental evidence that these role model effects exist.

Second, we contribute to the literature on gender differences in the willingness to compete mentioned above (Gneezy, Niederle, and Rustichini (2003) and Niederle and Vesterlund (2007)).[3] Gneezy, Niederle, and Rustichini (2003) conduct a laboratory experiment where subjects are asked to solve computerized mazes. They find that women and men perform similarly in a non-competitive environment, but only men increase their performance with increasing competitive pressure, which leads to a gender gap in performance in tournaments. Niederle and Vesterlund (2007) show that women are less likely to choose a tournament compensation scheme than men even when gender differences in performance do not exist. Our paper provides the first evidence that competitive female role models close the gender gap in tournament entry.

Finally, we contribute to the literature on women in top management positions (Adams and Kirchmaier (2016), Adams and Funk (2012)). Adams and Kirchmaier (2016) point out the low fraction of women in boardrooms. One reason for this observation may be that women shy away from competition. Thus, they may feel less comfortable to self-select into segments of the labor market that require competitive behavior and a high level of self-confidence. Our results suggest that female role models can mitigate this effect and encourage more women to enter competitive environments like climbing up the corporate ladder.

Our paper has important policy implications. Currently, a prominent form of intervention with the aim of decreasing gender gaps are gender quotas. However, the effectiveness of quotas is a much debated topic. Ahern and Dittmar (2012) show that

[3]For an overview see also Bertrand (2011).

the introduction of a gender quota for the board of directors in Norway led to a significantly negative announcement effect of affected firms.[4] The authors attribute their result partly to the fact that there were not enough sufficiently qualified women that firms could appoint to their boards after the quota became effective. At the same time, Bracha, Cohen, and Conell-Price (2018) find that a gender quota in a tournament based on a difficult math test decreased the performance of high ability women, possibly due to the quota intensifying stereotype threat. Our results suggest that there may be a positive *indirect* effect of gender quotas: If more women are appointed to boards, more female role models are available such that more other (highly qualified) women might eventually decide to enter a competitive career and aim for a top management position. Thus, while the immediate reaction to a gender quota is negative, it might still have positive consequences in the mid-term.

2 Data and summary statistics

2.1 Baseline experiment without role models

We base our experimental design on the design of Niederle and Vesterlund (2007), which consists of four rounds. In the first three rounds, subjects have to solve the same task of adding up as many sets of five two-digit numbers in three minutes as possible. Their compensation differs in each of these rounds. In Round 1, subjects were paid for their performance with a non-competitive piece-rate scheme. Specifically, subjects received 50 cents for each correct answer they gave in this round.

In Round 2, only the best performing subject with the highest number of correct answers in a group of four was paid and received $4 for each correct answer.

[4] The results of this paper have recently been criticized by Eckbo, Nygaard, and Thorburn (2016).

Niederle and Vesterlund (2007) designed the payment structure in their experiment so that a risk neutral subject with a 25 percent chance of winning the tournament would have the same payoff from the tournament as from the piece rate, i.e., subjects received $2 for each correct answer in the tournament. We increased the payment in the tournament in order to increase the incentives for choosing the tournament. Since subjects could not see each other during the experiment and also have no control over how we, as the experimenters, adhere to the way we calculate the bonus payments, the perceived risk of the tournament compared to the piece-rate was much higher than in a laboratory environment. Therefore, we needed to provide subjects with additional compensation to compensate for taking this risk.[5]

These first two rounds mainly serve to familiarize subjects with the task as well as the two different compensation schemes, that is, subjects do not have a choice on how they are compensated. These rounds are followed by Round 3, which is the main round of the experiment. In this round, subjects are asked which of the two compensation schemes they want to be applied to their performance. If they choose the piece-rate scheme, they again receive 50 cents for each correct answer independent of the other group members' performance. If they choose the tournament scheme, they receive $4 for each correct answer if they provided more correct answers in Round 3 than the other group members in Round 2. If they do not answer more problems correctly in this round than the best member in Round 2, then they do not receive any payment in Round 3. Niederle and Vesterlund (2007) suggest to determine payoffs in this round based on the relative performance regarding the other members in Round 2 because in this way the participants' choice does not influence the payoff of any other group member and it also does not depend on the tournament entry decisions of the other players.

[5] In a pre-test, barely anybody chose the tournament compensation scheme when we offered $2 for each correct answer to the best group member.

In the fourth round, subjects do not have to perform any calculations anymore and are only asked which compensation scheme they want to be applied to their past performance in Round 1. Thus, Round 4 differs from Round 3 because subjects no longer need to perform under a competition. This setup allows us to differentiate the actual act of performing under competition from other gender differences, for example in risk aversion and overconfidence, that may be also be causing the gender gap in tournament entry. Although the focus of our paper is on role model effects rather than the different channels contributing to the gender gap in tournament entry, for completeness and comparability with Niederle and Vesterlund (2007), we follow their approach and also include it in our experimental design.

In Niederle and Vesterlund (2007), subjects are not allowed to use a calculator to solve the problems. However, since our experiment is not conducted in a laboratory but each subject takes part in the experiment remotely, we have no possibility to enforce this rule. Since subjects have an incentive to break this rule to increase their performance-based compensation for participating in the experiment, we have to assume that some will. Forbidding a calculator without the means to enforce this rule would induce several unobservable factors that may impact the decision to compete or not compete and that may also be correlated with gender, mainly the propensity to cheat (Gino, Krupka, and Weber (2013)), and the trust put into the honesty of other subjects (Croson and Gneezy (2009)). Furthermore, allowing a calculator has the advantage that the task is a real effort task that does not require any prior mathematical skill.

The sets of two-digit numbers that subjects have to add up are randomly generated. Only one problem is presented at a time. As soon as a subject submits an answer a new set of numbers appears. Subjects are informed immediately whether their answer was correct or not. The total number of correct and incorrect answers are always shown in the upper part of the screen. Their final score for each round consists of the total number of correct answers given in three minutes. Hence, subjects always know their absolute performance in a task but they are not informed about their performance compared to the other members in their group. Each group consists of two men and two women who were randomly assigned to the groups.

After the experiment ended, subjects' performance-based compensation is determined based on only one of the four rounds that is randomly determined. Thus, subjects cannot use their decision in one round to hedge against the outcome of another round.

Additionally, we want to measure how confident subjects are in their performance and how this impacts their choice for competition. Whereas subjects are told their absolute performance throughout and after each task, they are not informed about their relative performance in their group. After completing all four rounds, we ask subjects to estimate their relative rank in Round 1 and Round 2. Subjects are paid $1 for each correct guess.

2.2 Role model conditions

To test our hypotheses, we use a 3 (female role model, male role model, no role model) x 2 (subject gender) between-subject design. According to Bandura (1986), an observer is more likely to imitate the behavior of a role model if it is perceived as similar and likeable. In addition, role models in our context need to be perceived as competitive since they are supposed to encourage competitive behavior.

In a pretest which is described in detail in Meier (2017), sixteen female and male individuals are evaluated based on a survey to find out whether they could potentially serve as role models for competitiveness. Role models' perceived competitiveness is measured using the four items on competitive motivation from the Motivational Trait Questionnaire from Heggestad and Kanfer (2000). Likeability was measured using three semantic differentials (likeable - unlikeable, agreeable - disagreeable, and pleasant - unpleasant). The role model questions were taken from Ragins (1999). Subjects gave their answer to all items on 7-point Likert scales.

Out of the sixteen potential role models considered in Meier (2017), we take the two male and female role models that were perceived as most competitive and displayed equal levels of likeability and role model potential. All of them are successfully working in competitive environments and are interviewed about their career path. In the videos, they all stress the need to engage in competitive behavior in order to be successful.

The role models are working in two different competitive fields, sports and business. One male and one female role model are famous Tennis players: Serena Williams and Roger Federer. Both of them had worldwide success for a number of years and became famous for winning an extraordinary number of the top tennis tournaments. The other two role models are successful business people: Marc Cuban and Nour Al Nuaimi. Marc Cuban, an American business man, became famous for his role as investor in "Shark Tank", an ABC reality television series. The show is about aspiring entrepreneur-contestants making business presentations and competing for funding. Nour Al Nuaimi is a private equity and venture capital investor, who has been interviewed on her career aspirations and the enjoyment of competing. Appendix B presents screenshots of the role model videos used in our experiment.

In addition, we chose a video of an Australian landscape for the neutral condition, which contained no human elements such as voices or persons. Since mental capacities or other influences can impact decisions, we wanted the three conditions (female role model, male role model, and no role model), to be as similar as possible. For this reason, we decided to also show subjects in the neutral condition a video instead of not showing them any video.

Before the addition task, subjects were randomly allocated to one of the 5 videos. Information on subjects' gender additionally allowed us to control for an even gender distribution across all treatments. Subjects were told that they would be shown a video and that they would subsequently have to answer questions about the person in the video. Each video lasted around 3 minutes. Subjects could not pause or rewind the video. After the video was finished, subjects were automatically forwarded to the next page. We then elicited answers to the same items as in the pretest as well as an item on the self-perceived similarity between role model and subject as a manipulation check.

We activate stereotype threat before subjects saw the role models following Marx and Ko (2012) by making them aware that they will be evaluated based on their performance and that they will receive feedback on this performance via a payment. Further, we elicit demographic information including subjects' gender before subjects watch the videos.

2.3 Experimental procedure

We conduct our experiment on Amazons Mechanical Turk (AMT) platform. We chose to run the experiment on this platform rather than in the laboratory, because the number of observations that researchers can elicit is much higher compared to a laboratory setting, with (in our case) 24 computers and thus 24 observations per session.

AMT is an online platform similar to a labor market, in which requestors can list tasks (called human intelligence tasks or HITs) along with a specified compensation. Individual "workers" can then elect to work on the task. They first click on a link to view a brief description of the task. This link can be tied to requirements that a worker has to fulfill (e.g. only US workers can open the link). After seeing the preview, workers can choose to accept the HIT, at which point the work is officially assigned to them and they can begin completing the task. Requestors can set a time frame in which the task has to be completed. Once a worker submits the result of her work, requestors decide about the quality of the work and how much the worker is paid. If the worker fulfilled the task properly, she will get the specified compensation, if she does not, requestors can neglect payment. Requestors can also pay a bonus if workers fulfilled a task extremely well.

Three sets of issues might introduce a bias in our AMT procedure. First, the AMT population is a group with access to the internet and knowledge of the existence of the AMT platform. Thus, the AMT population may not be representative of the average population, which also comprises a substantial fraction of older people. However, this critique might be even more severe in laboratory experiments where mostly university students participate.

Second, the AMT workers might not take their task seriously. We try to mitigate such concerns by selecting only workers with an approval rate above 95 percent, i.e., workers that have been rated to be highly reliable on previous tasks. Since these workers already received compensation for completing various tasks for other requesters, and since previous work performance is an important criterion for receiving future tasks, reputation should be important for them. Our results indicate that most workers took their task seriously as indicated by the time needed to complete the experiment (ranging from 16 minutes to 1 hour with an average of 28 minutes).

Third, AMT workers were not directly exposed to a tournament situation as they only obtained information about their group online but were not sitting with other group members in the same room (as would be the case in the laboratory). We think that this would only work against us, because the environment that AMT workers are in should be less intimidating for women, make stereotype threats less salient and thus might encourage more women to participate in a tournament compared to a laboratory environment.

We ran the experiment in three sessions in October and November 2016 and in May 2017. The first two sessions comprised approximately half of our sample. To receive a homogenous sample of AMT workers and to make our results comparable to Niederle and Vesterlund (2007), we required all workers to be located in the US and have an approval rate of at least 95 percent. We eliminate subjects' second participation if they took part in the experiment more than once. Our final sample consists of 847 subjects. After each session, we randomly assign each group to one of the payment rounds. Subjects were paid a participation fee of $5 and an additional performance-based fee averaging $10.89.

The description of the HIT said that workers could earn a bonus additional to their show-up fee depending on their performance in the second part of the HIT. Workers were then told to follow the link on the bottom of the page that would lead them to our study. Upon completion, they would receive a code that they should enter in AMT to receive credit for completing the HIT. This allowed us to verify that the worker had actually completed our experiment and allowed us to match workers' performance with their worker ID to pay them a performance-dependent bonus for the round which was randomly drawn in the end of the experiment. The description also required workers to be able to watch a video with sound on their computer. We neither mentioned that the goal of our experiment is to investigate gender differences in tournament entry in the recruitment process of workers nor in the description of the HIT. However, subjects were explicitly told that they were going to be allocated to mixed-gender teams of two women and two men, respectively.

All experiment instructions were provided to the subjects in written form on their screens. We programmed the experiment using Sosci Survey (Leiner (2015)) and used their server to collect the data.[6]

2.4 Summary statistics

Table 1, Panel A, shows the mean, median, standard deviation, 25th and 75th percentile of the variables collected from all subjects in our experiment, as well as the corresponding number of observations.

Subjects are on average between 30 and 34 years old and hold a high school diploma or a similar secondary school certificate. When asked how important knowledge in math is to them, most subjects indicated that it is important for them to be good in math (average score of 3.77 on a scale ranging from zero (not important) to six (very important).

With respect to math performance in all conditions, we find that the average number of correctly answered addition problems (math score) is 11.69, 13.07, and 13.73 in Round 1 (piece rate compensation scheme), Round 2 (tournament compensation scheme), and Round 3 (subjects' choice of compensation scheme), respectively. The increase in performance is likely to be due to learning effects and an increase in performance caused by the tournament conditions in Round 2 and Round 3. Subjects improved their score from Round 1 to Round 2 by 1.38 correctly answered problems

[6]For more information on this tool, see https://www.soscisurvey.de/.

on average. There also is an increase in performance from Round 2 to Round 3, but it is smaller in magnitude.

Subjects took an average of 9.80 seconds to decide whether they wanted the piece-rate or the tournament payment scheme to be applied to their performance in Round 3. The average time until a decision was made decreased to 6.45 seconds in round 4. Most subjects chose the piece-rate in both rounds. 22 percent of subjects chose the tournament in Round 3 (choice 1) and 15 percent chose the tournament in round 4 (choice 2).

To elicit the presence of stereotype threats, we asked subjects whether they agree to the statement that men are better at solving math-related problems than women. About 15 percent of subjects agreed or strongly agreed to this statement.

In Panel B of Table 1, we compare the mean values of female and male subjects across all conditions (i.e., neutral, female role model, and male role model). On average, female subjects are significantly older than male subjects. There are no significant gender differences in education or subjects' perceived importance of math.

With respect to performance, we find no significant gender difference in the number of correctly solved math problems in Round 1 (piece rate compensation scheme). However, male subjects perform slightly better than female subjects in Round 2 (tournament compensation scheme). This finding is in line with Gneezy, Niederle, and Rustichini (2003), who find that men increase their performance under competition more than women. This effect becomes larger in Round 3 (choice of compensation scheme), since male subjects chose the tournament more frequently and were thus more likely to perform under competition.

Regarding tournament entry, we find that male subjects choose the competitive compensation scheme significantly more often than women in both, Round 3 and round 4 (corresponding to choice 1 and choice 2). 28 (20) percent of male subjects enter the tournament in Round 3 (round 4), whereas only 17 (11) percent of female subjects enter the tournament.

Finally, we also find a pronounced gender difference with respect to our stereotype threat variable. Male subjects seem to have more negative stereotypes with respect to the math ability of women than female subjects. Male subjects' agreement with the statement that "men are better at solving math related problems" is on average 1.95, whereas female subjects only agreed with a score of 1.39.[7] This difference is highly statistically significant (t-stat: 4.68). Even if women do not agree to the statement themselves, they may be well aware that men hold this view, which would suggest that they are under stereotype threat.

3 Tournament entry without role models

We first examine whether our experimental design delivers the same baseline result as in Niederle and Vesterlund (2007) in our neutral condition, where no role models are introduced.

To make conditions comparable with respect to the presentation of visual stimuli, subjects in this condition are also shown a video. However, the video contained no human elements. Instead, the video shows footage from the landscape of and around Uluru in Australia, accompanied by atmospheric instrumental music without any singing. We chose a video that would not be activating in any way as this has been shown to change risk behavior (Andrade, Odean, and Lin (2015)).[8]

[7] The item was adapted from Marx, Monroe, Cole, and Gilbert (2013). The overall low rating could be due to a bias towards socially desirable answers.

[8] The video was created by Selmesfilm. It can be accessed here: https://www.youtube.com/watch?v=FJPwPqvZaUI.

Similar to Niederle and Vesterlund (2007), we find no gender difference in performance under the piece rate or under the tournament condition among AMT workers, too. The average number of problems solved in the piece-rate (tournament) compensation scheme in the neutral condition is 11.17 (12.37) for female subjects and 10.90 (12.66) for male subjects. Two-sided t-tests reveal that these differences are not statistically significant.

Next, we analyze subjects' choice in Round 3, i.e., whether to get paid according to a piece rate scheme or a tournament scheme, respectively. Table 2, Panel A, shows the frequency of men and women that entered the tournament in the neutral condition without role models. Compared to Niederle and Vesterlund (2007), tournament participation rates are lower for both, female and male subjects. While Niederle and Vesterlund (2007) find that around two thirds of men chose the tournament, only 25 percent of men in our experiment choose the tournament. Also, only 14 percent of women choose the tournament compared to 35 percent in Niederle and Vesterlund (2007). Nevertheless, we find a significant gender difference in tournament entry. Among subjects that chose the tournament, there are 33 percent less women. A Pearson's Chi-squared test confirms that this difference is also statistically significant (p-value = 0.087).

The overall lower preference for competition may be due to two differences in our experimental procedure working together: First, compared to a laboratory setting, subjects could not see the other group members and are just told that they are competing against three other subjects in a gender balanced group, thus creating a greater uncertainty with respect to whom exactly the subject is competing against. In addition, there is probably less trust towards the experimenter since it is almost impossible for participants to control whether we paid out the correct amount in the tournament. In the piece rate, by contrast, subjects can calculate their bonus on their own if they remember their number of correct answers in a specific round.

In Table 2, Panel B, we compare the past performance of men and women by choice of compensation scheme. We find similar results as Niederle and Vesterlund (2007). Male subjects who performed better in the piece rate and tournament conditions before are significantly more likely to enter the tournament. Also, male subjects who increased their performance more strongly from Round 1 to Round 2 are more likely to choose the tournament in Round 3. For female subjects, performance in the first two conditions is not significantly related to the subsequent decision on tournament entry. Thus, in contrast to male subjects, the decision to enter the tournament must depend on factors other than math performance for female subjects.

In Figure 1, we compare the propensity to enter the tournament for female and male subjects over different performance quartiles. Since the optimal decision depends on subjects' estimates of their future performance in Round 3, we show the tournament entry decision conditional on subjects' actual performance in Round 3 (right panel) as well as their past performance in Round 2 (left panel). With the exception of the third performance quartile in Round 2, where the percentage of male and female subjects is approximately equal, male subjects are always more likely to choose the tournament compensation scheme in each performance quartile in Round 2 and Round 3. Even female subjects in the highest performance quartile enter the tournament less frequently than the average male subject. With respect to subjects' performance in Round 2, the gender gap actually widens for the two highest performance quartiles.

Finally, we examine gender differences in tournament entry based on a probit model including various control variables. Marginal effects are presented in Panel C of

Table 2. In these regressions, the dependent variable is equal to one if a subject chooses the tournament compensation scheme in Round 3, and zero otherwise. Column (1) shows the baseline result without any controls. In column (2) we add performance controls, i.e., subjects' performance in the piece rate compensation scheme (Round 1) and their performance improvement between rounds 1 and 2.[9] In column (3), we also add controls for subjects' education and for subjects' perceived importance of math.

Confirming our univariate results from Panel A in Table 2, results in column (1) show that female subjects are 10.8% less likely to enter the tournament than male subjects. The difference is statistically significant at the 10% level. Controlling for subjects' performance in column (2) does not alter this finding. Female subjects are still 11% less likely to enter the tournament than male subjects. The effect gets larger in both, economic and statistical terms, once we control for differences in education and the perceived importance of math (column (3)). In this specification, female subjects are 16.4% less likely to enter the tournament than male subjects. This difference is statistically significant at the 5% level.

4 The impact of role models on tournament entry

4.1 Validation of role model choice

For our treatments to work, the individuals we present in our videos need to be perceived as role models for competitiveness. Furthermore, the literature on role models

[9]Niederle and Vesterlund (2007) include tournament performance rather than piece rate performance as control variable. Including either of these variables does not change our main result. However, we think that including piece rate performance is the better choice from a conceptional point of view, as it more accurately captures math performance and is less likely to be influenced by gender differences in performance due to the tournament condition rather than math skills.

argues that their behavior is more likely to be imitated if they are perceived as likeable and similar to the observer (Bandura (1986). We therefore first examine whether female and male subjects in our experiment perceived the four role models as being competitive, likeable and similar to themselves, thereby validating that we have chosen useful role models for the purpose of our experiment. Table 3 displays the mean ratings for competitiveness, likeability, and similarity across subject gender and role model gender.

Results in column (1) show that the average competitiveness score across all role model videos is significantly higher than the scales midpoint (3.00) with an average score of 4.62 across all videos. With respect to role models' gender, we find that male role models are perceived as slightly more competitive with a mean score of 4.73 compared to a mean score of 4.51 for female role models ($p = 0.004$). This difference should work against finding a reduction in the gender gap in tournament entry for male vs. female role model videos, as male role models may encourage competitive behavior more strongly than their female counterparts. Moreover, the different perception seems to stem mostly from male subjects ($p = 0.000$), whereas female subjects did not perceive the competitiveness of female and male role models differently ($p = 0.943$). Rather, they perceived both, male and female role models, as (equally) competitive.

In column (2), we compare the likeability ratings of our role models. Results show that all role models are perceived to be highly likeable. Mean likeability scores for each role model are significantly larger than the midpoint of the scale (3.00), ranging from 4.40 to 5.08. This finding suggests that they are more likely to be imitated by observers than if they had received low likeability ratings.

Marx and Ko (2012) find that high similarity of the role model to its observer increases its effectiveness. Thus, we also examine whether the female and male role models we introduce are perceived to be similar to the subjects in our sample. Results are presented in column (3) of Table 3. As expected, male subjects perceive male role models to be more similar to them than female roles models. Thus, male role models should have a stronger impact on male subjects' behavior than female role models. In contrast, female subjects perceive both, female and male role models, to be significantly less similar to them compared to the midpoint of the scale. This may be related to the fact that all role models display competitive behavior and could lead to a weaker impact of role models on female subjects' behavior.

Overall, our manipulation checks confirm that the role models we use are perceived as highly competitive and likeable by subjects and should thus encourage competitive behavior. Perceived similarity of role models to subjects is rather weak, which should rather work against us and may result in generally weaker role model effects, particularly for female subjects, in our subsequent analysis.

4.2 Tournament entry in role model conditions

We now turn to an analysis of our main research question whether exposing female subjects to competitive role models changes their propensity to participate in tournaments. To do so, we first compare female subjects' propensity to enter the tournament when they were shown a male vs. a female role model, respectively. Table 4 reports marginal effects from a probit model where the dependent variable is equal to one if a subject chooses the tournament compensation scheme in Round 3, and zero otherwise. We use the same model specifications as in Panel C of Table 2.

Panel A of Table 4 displays tournament entry decisions for subjects who are shown a male role model (i.e., Marc Cuban or Roger Federer). In this condition, the coefficient on the gender difference in tournament entry becomes highly statistically significant. Specifically, female subjects are 14.2% (column (2)) to 20.6% (column (3)) less likely to enter the tournament than male subjects. The gender difference in tournament entry is significant at the 1% level for all specifications. Thus, compared to the neutral condition without role models (see Panel C of Table 2), male role models increase the gender gap in tournament entry even further. The reason may be that male role models trigger stereotype threats for female subjects, as male role models are perceived as very competitive (see Table 3), which may intimidate female subjects and further decrease their willingness to enter a tournament.

Regarding the impact of the control variables, we also find that subjects' piece rate performance significantly increases the propensity to enter the tournament (column (2)).

Panel B of Table 4, presents results for subjects who are shown a female role model (i.e., Nour Al Nuaimi or Serena Williams). In this condition, the gender gap in the propensity to enter the tournament disappears. Marginal effects suggest that female subjects are 5.4% (column (3)) to 7.5% (column (1)) less likely to enter the tournament than male subjects. However, this difference is not statistically significant anymore. Presenting a competitive female role model to female subjects thus indeed seems to encourage more women to enter the tournament.

After establishing our main result that female role models lead to a disappearance of any significant gender gap in tournament entries, we now analyze decisions in Round 4 in which subjects do not have to perform calculations anymore, but simply decide whether they would like a piece-rate or a tournament compensation scheme applied to their performance in Round 1. Niederle and Vesterlund (2007) use this round to examine whether gender differences in tournament entry are due to lower preferences of female subjects to perform in a competition, or due to other explanations such as gender differences in risk aversion and overconfidence. To differentiate between these explanations, they include two additional control variables, i.e. subjects' beliefs about their own performance and their choice of compensation scheme in Round 4. Following this approach, we re-run our regressions of Table 4 and also include these additional control variables.

Results are reported in a table in Appendix C. In Panel A, we examine gender differences in tournament entry in the male role model condition, while Panel B presents results for the female role model condition. When adding subjects' beliefs regarding their tournament performance as additional control (column (1)), we find that in both conditions, subjects, who believe that they only achieved a low performance rank in the previous tournament (Round 2), are significantly less likely to enter the tournament in Round 3. In column (2), we add subjects' choice in Round 4, i.e., a dummy variable equal to one if subjects choose the tournament compensation scheme to be applied to their performance in Round 1, and zero otherwise. We find that subjects are significantly more likely to choose the tournament in Round 3 if they also choose it in Round 4. This result obtains if we include both control variables in the same regression (column (3)), while the impact of subjects' beliefs on tournament entry become insignificant in this specification.

Most importantly, adding subjects' beliefs and their choice in Round 4 as additional control variables does not alter our main result: Coefficients on the female subject dummy in Panels A and B still show that a male role model further widens the gender gap in tournament entry, while the female role model decreases the gender gap in tournament entry, resulting in an insignificant female subject dummy in

Panel B. Assuming that beliefs about subjective performance and subjects' choice in Round 4 sufficiently capture gender differences in risk aversion and overconfidence, these results suggest that role models have a direct impact on subjects' preferences to perform in a competition.

Taken together, results in this section show that the effect of female role models on the gender gap in tournament entry is opposite to the effect of male role models: While the former leads to an insignificant gender gap in tournament entry, the latter widens the gender gap and further reduces women's willingness to compete against others. The reason for this difference may be that male role models trigger stereotype threats, while female role models lead to an encouragement effect. Encouragement effects would be particularly desirable if especially well-performing women became more likely to compete. We will further examine the interaction of role model effects, stereotype threats, and performance in our subsequent analysis.

5 Which female subjects react most?

We now turn to analyze whether competitive behavior of certain groups of women is particularly affected by the presentation of female role models. We first analyze the effect of role models on high vs. low performing female subjects in Section 5.1. Then we examine whether female role models have a stronger impact on women subject to stronger stereotype threats in Section 5.2.

5.1 Impact of role models on high vs. low performing women

Efficiency losses can occur both, when high performing subjects refrain from entering the tournament even though they would have good chances of winning, but also when low performing subjects enter the tournament even though they have little chances of winning. Thus, it is important to note that increased tournament entry rates are not necessarily desirable per se. Role models only promote optimal behavior if they encourage high performing female subjects who shy away from competition to enter the tournament more frequently, while at the same time not encouraging low performing women for whom it would be suboptimal to enter the tournament.

Previous literature in psychology suggests that high and low performing women should indeed react differently to female role models. According to Marx and Roman (2002), women with a high math competence may be most subject to stereotype threat and thus profit most from role models, which buffer this threat. In addition, subtyping may impede the impact of role models on women with low prior performance in math. According to this concept, members of a group may view an individual that disconfirms stereotypes of that group as being an exception from the rule and place them in a separate category (Richards and Hewstone (2001), Ziegler and Stoeger (2008)). Moreover, a lack of perceived attainability may lead subjects to become intimidated by dissimilar role models such that they cannot benefit as much from the role model as subjects with higher perceived attainability (Marx and Ko (2012), Lockwood and Kunda (1997)). Therefore, we test the conjecture that the positive influence of a female role model on the likelihood of female subjects to enter the tournament will be concentrated on women who perform better in the math task.

To test this hypothesis, we first plot the tournament entry decisions over different performance quartiles from Round 2 and Round 3 for female subjects who saw a female and a male role model separately in Figure 2.

Visual inspection shows that female role models have the strongest impact on the best performing women, i.e., those belonging to the top performance quartile. Compared to the best performing women who observed a male role model, they choose the tournament more frequently: while about 40.0% of the best performing women being exposed to a female role model enter the tournament, only about 14.3% of them do so after being exposed to a male role model. In contrast, 42.5% of the best male subjects enter the tournament after seeing a male role model and this number drops to 23.3% after seeing a female role model. The decision of female subjects to enter the tournament is less influenced by role models in the other three performance quartiles.

Shying away from competition is very costly for the best performing women in our experiment. The top quartile of women earn on average 8.67 under the piece-rate payment scheme. If they win the tournament, they earn on average 50.34 and their probability of winning the tournament is 65% , i.e., the expected remuneration when these women select the tournament payment scheme is 32.72 or nearly four times as large than under the piece-rate condition. Of course, the standard deviation of payoffs in the tournament condition is higher. However, to rationalize a decision of the best performing females to choose the piece-rate would obviously require implausibly high levels of risk aversion.

In the next step, we turn to a more formal investigation and again run probit regressions with tournament entry as the dependent variable. Probit regressions are run separately for two groups of subjects, those in the highest and lowest performance quartile in the sample based on their Round 1 performance. Results are reported in Table 5. In column (1), the sample is restricted to female subjects only, while column (2) presents results for male subjects only. Results in column (3) are based on both, female and male subjects.

Results in Panel A (B) of Table 5 are based on subjects belonging to the highest (lowest) performance quartile only. The variables of interest are Female role model, Male role model, and the interaction term between Female role model and Female subject. Female (male) role model is a dummy variable with the value of one if the subject saw a female (male) role model, and zero otherwise. The interaction term Female role model × Female subject is one if the subject is a women and she saw a female role model, and zero otherwise. We use the same set of control variables as before. Columns (1) and (2) of both panels report marginal effects estimated for discrete changes from the baseline for indicator variables and at the means for all other variables. Column (3) shows the coefficient estimates because the inclusion of the interaction term does not allow to compute marginal effects.

Results in column (1) of Panel A show that observing a female role model has a significantly positive effect on female subjects' propensity to enter the tournament in the highest performance quartile. Specifically, female subjects are 16.8 percentage points more likely to enter the tournament if they were shown a female role model. This difference is significant at the 10% level.

By contrast, results in column (2) of Panel A show that being shown a male role model does not impact the tournament entry decision for the best performing men. This result is similar to Carrell, Page, and West (2010), who show that professor gender does not impact male students' performance or selection of future courses. One potential reason for male students not reacting to role models may be that they do not face a stereotype threat or are not exposed to social norms discouraging them from competing that could be buffered by a role model.[10] When pooling together both male and female subjects, we find that female subjects are again significantly less likely to enter the tournament. However, this effect is significantly reduced if a female subject is shown a female role model. In economic terms, only 17.6% of highly performing female subjects who were not shown a female role model enter the tournament, while 31.2% of highly performing female subjects who saw a female role model enter the tournament.

Results in Panel B show the coefficient estimates for the subjects in the lowest performance quartile. For this group, there are no significant role model effects, neither for the female nor for the male subjects.

In summary, we find support for our conjecture that the positive influence of female role models on women's likelihood to compete is concentrated on the highest performing women. Since these women are most likely to win the tournament, it may be that they are only hindered by low self-confidence or stereotype threat and that role models buffer these effects. In line with this argument, we earlier found that tournament performance predicts entry for men but not for women in Table 2. https://v2.overleaf.com/project/5922c3c6e3dc59486eae2c6b

[10]In unreported results, we also perform a probit regression for men with a dummy for the female role model. The coefficient is insignificant. Thus, the positive effect of a female role model does not seem to stem from idiosyncracies of the female role models. Also female role models do not seem to decrease competitive behavior in men.

5.2 Impact of role models on women with high vs. low stereotype threat

According to Steele and Aronson (1995), an individual exposed to stereotype threat feels at risk of confirming a negative preconception about his or her group. As a result, stereotype threats lead to lower self-confidence and performance expectations (Stangor, Carr, and Kiang (1998)). Results in Table 1 suggest that male subjects are more likely to hold the view that they are better in solving math related tasks than their female counterparts. This view may trigger a stereotype threat among female subjects that reinforces their reluctance to compete against men in a math-related task. Female role models may mitigate these effects as they pose examples of success in a stereotyped domain which may increase female subjects' willingness to participate in the tournament themselves. To test this conjecture, we first examine whether stereotype threats regarding gender differences in math competence indeed influence subjects' beliefs about their performance. In the experiment, we ask subjects to guess their performance rank in the piece rate compensation scheme based on their own subjective judgment ranging from 1 (highest performance rank) to 4 (lowest performance rank). Given that there are no gender differences in piece rate performance (see Table 1) and there are no tournament considerations in Round 1, there should be no difference in average guessed performance ranks between female and male subjects. However, stereotype threat may lead to female subjects estimating their performance to be lower than male subjects.

In Table 6, we regress subjects' guessed piece rate performance on our measure of stereotype threat, subjects' true piece rate performance and indicator variables for the different role model conditions. Results are based on female subjects only.[11]

[11] We do not find any significant role model effects for male subjects.

They show that female subjects provide higher subjective performance rankings for themselves after they have observed a female role model. In economic terms, the probability that a female subjects sorts herself in the highest performance rank is 5.4% higher after she has been shown a female role model compared to the neutral condition. We do not find a similar effect for female subjects who were shown a male role model. Results in column (2) suggest that the stereotype threat leads to lower guessed performance ranks for female subjects, however, the coefficient is not statistically significant. In column (3), however, female subjects who agree to the statement that men are better in solving math problems (stereotype threat) rank themselves significantly lower in terms of guessed performance than female subjects who do not agree to this statement to the same extent. In all specifications, realized performance and guessed performance ranks are significantly related, indicating that female subjects who performed well in the addition task predicted their performance rank to be higher.

In the next step, we examine whether female role models not only mitigate the negative impact of stereotype threat on female subjects' estimation of their own performance, but whether there is also a positive impact on tournament entry probabilities. We first sort female subjects into high and low stereotype threat groups according to the median of the stereotype threat variable.[12] Results are reported in Table 7.

Comparing coefficients in columns (1) and (2) shows that exposure to a female role model has a particularly strong and positive impact on female subjects that feel strong stereotype pressure. In this group, female subjects who are shown a female role model are 10.9% more likely to enter the tournament than female subjects who

[12] Stereotype threat is measured on a scale from zero to six using the item: "Men are better at solving math related problems." Therefore, we use the median as a cutoff to define high vs. low stereotype threat.

are not shown a female role model. There is no significant difference in tournament entry between female subjects in the low stereotype threat group. These results are confirmed if we pool all observations and include an interaction term between the female role model variable and the stereotype threat variable (column (3)). We find that female subjects who feel a strong stereotype threat are 14.1% less likely to enter the tournament. However, this result is mitigated if these women are shown a female role model, as indicated by the positive and significant interaction term.

Taken together, we find evidence that female role models mitigate the negative impact of stereotype threat on women's self-confidence regarding their relative performance within a group, as well as their reluctance to enter a tournament and compete against other.

6 Robustness checks

We conduct several additional tests to check whether our main result is robust to alternative model specifications. Most importantly, rather than estimating regressions for each treatment separately, we now examine the full sample and use interaction terms to check whether gender differences in the willingness to compete are different for the female role model treatment compared to the neutral and male role model treatment.

In Table 8, we regress the tournament decision on a dummy for being female, a dummy for the female role model treatment, and their interaction term. We again control for subjects' piece rate performance, performance increase, education, and importance of math. To show that our results are not driven by differences in subjects' confidence in their relative performance, we additionally control for subjects' guessed tournament rank. We also control for subjects' degree of reflection on the decision, proxied by the time they took to make their tournament entry decision.

Results in column (1) show that the interaction of *Female subject* and *Female role model* is positive and significant at the 5% level. Holding all other variables at their mean, the predicted probability of choosing the tournament almost doubles from 12.6% to 20.3% for a woman who saw a female role model compared to a woman who saw either the neutral video or a male role model.

In column (2), we examine whether subjects' choice depends on their tournament performance in task 2, rather than their piece-rate performance in task 1, which we use as a control variable in our main regressions. Our results are not affected. Similar to the results in Table 5, the baseline effect becomes more pronounced if we run our baseline regression only for subjects performing highest under the tournament payment scheme. With a predicted probability of 68.4%, women exposed to a female role model are now 4 times more likely to choose the tournament than women exposed to a neutral video or a male role model (15.4%).

In column (3), we check whether our results are indeed due to role model exposure by looking at the subset of subjects who indicated that the individual shown to them in the video would be a suitable role model to them. Specifically, we conduct a median split on the variable "role model", which we constructed from 4 items by Ragins (1999). Size and significance of the interaction slightly increases compared to column (1). However, the change in the predicted probabilities is almost the same. This suggests that the conscious level of identification with a role model is not greatly relevant for the effect of a female role model on women's willingness to compete.

7 Conclusion

There is broad evidence in the literature that, compared to men, women are less confident in their own abilities and shy away from competing with others. This effect has been suggested as one possible reason why there are only few women in top management positions or other comparable powerful positions. Many efforts have been undertaken by governments to increase the fraction of women in these positions with the aim of establishing gender equality.

In this paper, we suggest a new mechanism that help to encourage women to self-select in competitive environments and trust more in their own abilities: Exposing them to female role models, i.e., successful women who express their preference for competition and their aspiration to belong to the best. These role models encourage their female observers to view competition more positively and eventually self-select into competitive environments - a necessary condition to climb up the career ladder and reach top management positions. The effects we document are particularly strong among the best performing women, for whom it is most beneficial to enter competition. At the same time, female role models reduce the negative influence of gender stereotypes on women's self-confidence.

In contrast, male role models seem to reinforce gender stereotypes and increase the gender gap in tournament entry. In our experiment, women who were exposed to competitive male role models were even less likely to enter a subsequent tournament compared to women who did not see a role model at all.

Our findings suggest a positive impact of gender quotas on aspiring women by providing them with role models to follow. However, gender quotas would need to be constructed very carefully to make sure that no underqualified women are hired for or promoted to top positions and then eventually underperform. This would be interpreted as a negative role model and might discourage other women from entering career competitions in the first place. In situations where gender quotas could lead to insufficient incentives or increased stereotype threat to the most qualified women, prominently presenting role models could be an alternative.

References

Adams, R., and T. Kirchmaier, 2016, "Women on Boards in Finance and STEM Industries," *American Economic Review*, 106, 277–281.

Adams, R. B., and P. Funk, 2012, "Beyond the Glass Ceiling: Does Gender Matter?," *Management Science*, 58, 219–235.

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

Andrade, E. B., T. Odean, and S. Lin, 2015, "Bubbling with Excitement: An Experiment," *Review of Finance*, 20(2), 447–466.

Bandura, A., 1986, "Chapter 2 Observational Learning," in *Social Foundations of thought and action: A Social Cognitive Theory*. Englewood Cliffs, Englewood Cliffs, NJ, US, pp. 47–97.

Beaman, L., E. Duflo, R. Pande, and P. Topalova, 2012, "Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India," *Science*, 335(6068), 582–586.

Bertrand, M., 2011, "New Perspectives on Gender," in *Handbook of Labor Economics*, ed. by David Card and Orley Ashenfelter. Elsevier, Amsterdam; London, vol. Volume 4, Part B, pp. 1543–1590.

Bettinger, E., and B. T. Long, 2005, "Do Faculty Members Serve as Role Models? The Impact of Faculty Gender on Female Students," *American Economic Review*, 95, 152–157.

Bhana, D., 2005, ""I'm the best in maths. Boys rule, girls drool." Masculinities, mathematics, and primary schooling," *Perspectives in Education*, 23, 1–10.

Bracha, A., A. Cohen, and L. Conell-Price, 2018, "The heterogeneous effect of affirmative action on performance," *Journal of Economic Behavior & Organization*.

Bursztyn, L., G. Egorov, and S. Fiorin, 2017, "From Extreme to Mainstream: How Social Norms Unravel," Working Paper.

Bursztyn, L., T. Fujiwara, and A. Pallais, 2017, "'Acting Wife': Marriage Market Incentives and Labor Market Investments," Working Paper.

Cadsby, C. B., M. Servátka, and F. Song, 2013, "How Competitive Are Female Professionals? A Tale of Identity Conflict," *Journal of Economic Behavior & Organization*, 92, 284–303.

Carrell, S. E., M. E. Page, and J. E. West, 2010, "Sex and science: How professor gender perpetuates the gender gap," *The Quarterly Journal of Economics*, 125(3), 1101–1144.

Croson, R., and U. Gneezy, 2009, "Gender Differences in Preferences," *Journal of Economic Literature*, 47(2), 448–474.

Eckbo, B. E., K. Nygaard, and K. S. Thorburn, 2016, "Does Gender-Balancing the Board Reduce Firm Value?," CEPR Discussion Paper No. DP11176.

Else-Quest, N. M., J. S. Hyde, and M. C. Linn, 2010, "Cross-National Patterns of Gender Differences in Mathematics: A Meta-Analysis," *Psychological Bulletin*, 136, 103–127.

Fennema, E., P. L. Peterson, T. P. Carpenter, and C. Lubinski, 1990, "Teachers' attributions and beliefs about girls, boys, and mathematics.," *Educational Studies in Mathematics*, 21, 55–69.

Fisman, R., S. S. Iyengar, E. Kamenica, and I. Simonson, 2006, "Gender Differences in Mate Selection: Evidence from a Speed Dating Experiment," *Quarterly Journal of Economics*, 121, 673–697.

Flory, J. A., A. Leibbrandt, and J. A. List, 2015, "Do Competitive Workplaces Deter Female Workers? A Large-Scale Natural Field Experiment on Job Entry Decisions," *The Review of Economic Studies*, 82(1), 122–155.

Gino, F., E. L. Krupka, and R. A. Weber, 2013, "License to Cheat: Voluntary Regulation and Ethical Behavior," *Management Science*, 59(10), 2187–2203.

Gneezy, U., K. L. Leonard, and J. A. List, 2009, "Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society," *Econometrica*, 77(5), 1637–1664.

Gneezy, U., M. Niederle, and A. Rustichini, 2003, "Performance in competitive environments: gender differences," *Quarterly Journal of Economics*, 118, 1049–1074.

Günther, C., N. A. Ekinci, C. Schwieren, and M. Strobel, 2010, "Women Can't jump?—An Experiment on Competitive Attitudes and Stereotype Threat," *Journal of Economic Behavior & Organization*, 75(3), 395–401.

Heggestad, E. D., and R. Kanfer, 2000, "Individual Differences in Trait Motivation: Development of the Motivational Trait Questionnaire," *International Journal of Educational Research*, 33(7–8), 751–776.

Leiner, D. J., 2015, "SoSci Survey (Version 2.6.00-i)," http://www.soscisurvey.com.

Lockwood, P., and Z. Kunda, 1997, "Superstars and me: Predicting the impact of role models on the self.," *Journal of Personality and Social Psychology*, 73(1), 91 – 103.

Marx, D. M., and S. J. Ko, 2012, "Superstars "like" Me: The Effect of Role Model Similarity on Performance under Threat," *European Journal of Social Psychology*, 42(7), 807–812.

Marx, D. M., S. J. Ko, and R. A. Friedman, 2009, "The "Obama Effect": How a Salient Role Model Reduces Race-Based Performance Differences," *Journal of Experimental Social Psychology*, 45(4), 953–956.

Marx, D. M., A. H. Monroe, C. E. Cole, and P. N. Gilbert, 2013, "No Doubt About It: When Doubtful Role Models Undermine Men's and Women's Math Performance Under Threat," *Journal of Social Psychology*, 153(5), 542–559.

Marx, D. M., and J. S. Roman, 2002, "Female Role Models: Protecting Women's Math Test Performance," *Personality and Social Psychology Bulletin*, 28(9), 1183–1193.

Meier, K., 2017, "Women need female role models," Working Paper, University of Mannheim.

Niederle, M., 2015, "Gender," in *Handbook of Experimental Economics*, ed. by J. Kagel, and A. E. Roth. Princeton University Press, Princeton, NJ, vol. 2, pp. 481–553.

Niederle, M., and L. Vesterlund, 2007, "Do Women Shy Away from Competition?," *Quarterly Journal of Economics*, 122, 1067–1101.

——— , 2011, "Gender and Competition," *Annu. Rev. Econ.*, 3(1), 601–630.

Ragins, B. R., 1999, “Mentor Funtions and Outcomes: A Comparison of Men and Women in Formal and Informal Mentoring Relationships,” *Journal of Applied Psychology*, 84(4), 529–550.

Richards, Z., and M. Hewstone, 2001, “Subtyping and Subgrouping: Processes for the Prevention and Promotion of Stereotype Change,” *Personality & Social Psychology Review (Lawrence Erlbaum Associates)*, 5(1), 52–73.

Stangor, C., C. Carr, and L. Kiang, 1998, “Activating stereotypes undermines task performance expectations,” *Journal of Personality and Social Psychology*, 75(5), 1191–1197.

Steele, C. M., and J. Aronson, 1995, “Stereotype Threat and the Intellectual Test Performance of African Americans,” *Journal of Personality & Social Psychology*, 69(5), 797–811.

Stout, J. G., N. Dasgupta, M. Hunsinger, and M. A. McManus, 2011, “STEMing the Tide: Using Ingroup Experts to Inoculate Women’s Self-Concept in Science, Technology, Engineering, and Mathematics (STEM),” *Journal of Personality & Social Psychology*, 100(2), 255–270.

Ziegler, A., and H. Stoeger, 2008, “Effects of Role Models from Films on Short-Term Ratings of Intent, Interest, and Self-Assessment of Ability by High School Youth: A Study of Gender-Stereotyped Academic Subjects,” *Psychological Reports*, 102(2), 509–531.

Appendix

A Variable description

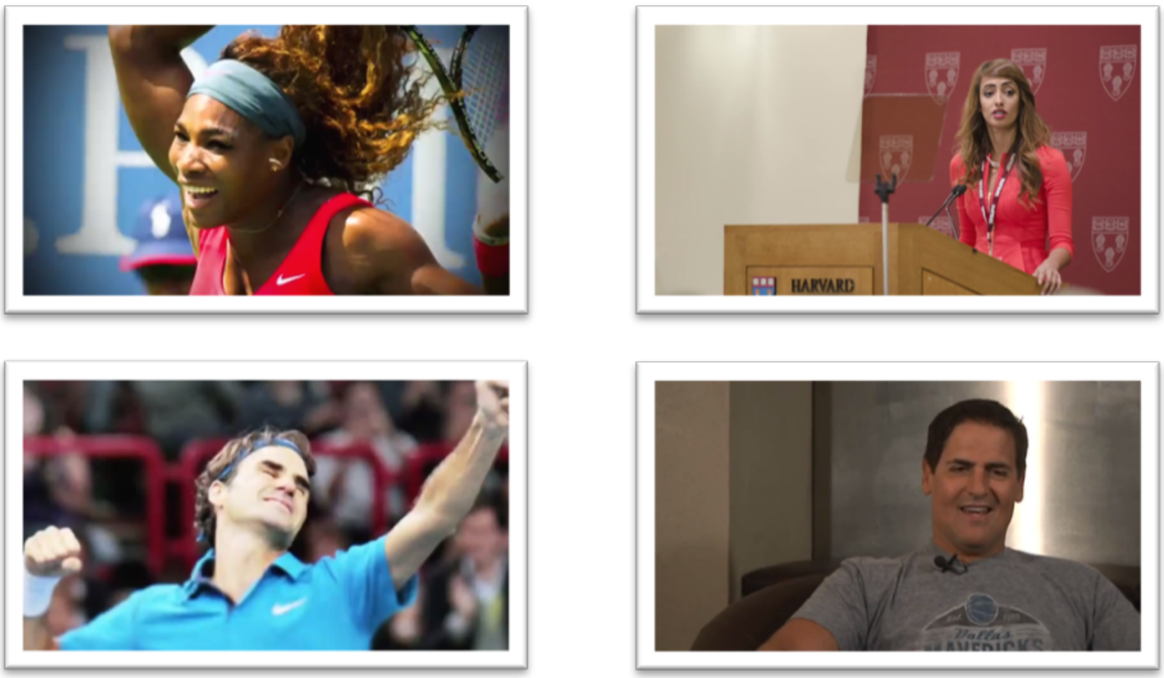
This table briefly defines the main variables used in the empirical analysis. All data are collected from amazon mechanical turk (AMT).

Variable name	Description
Subject age	An ordinary variable describing a subject's age, spanning from 0 = "younger than 15 years old" to 11 = "65 years or older" in steps of 15 years
Subject education	An ordinary variable describing a subject's highest educational achievement, spanning from 0 = "Left school with no qualifications" to 8 = "Vocational university / university of applied sciences / university degree"
Importance of math	A subject's degree of agreement to the statement "It is important for me to be good at math.", where 0 = "Very Untrue" and 6 = "Very True"
Math score 1	Number of correctly solved addition problems in Round 1
Math score 2	Number of correctly solved addition problems in Round 2
Math score 3	Number of correctly solved addition problems in Round 3
Performance difference 2−1	The difference between the number of correctly solved addition problems in Round 1 and Round 2
Time for choice 1	Time spent on the decision of the payment scheme for Round 3 measured as the time (in seconds) the subject spends on the page he or she needs to choose the payment scheme

Time for choice 2	Time spent on the decision of the payment scheme for Round 4 measured as the time (in seconds) the subject spends on the page he or she needs to choose the payment scheme
Choice 1	A dummy variable that takes on the value of 1 if a subject chooses the tournament payment scheme in Round 3 and 0 otherwise
Choice 2	A dummy variable that takes on the value of 1 if a subject chooses the tournament payment scheme in Round 4 and 0 otherwise
Stereotype threat	A subject's degree of agreement to the statement "Men are more skilled at solving math-related problems.", where 0 = "Very Untrue" and 6 = "Very True"
Role model	A subject's average response on a 7-point Likert scale to 4 items from Ragins (1999): "I think X's behavior is worth striving for," "X is someone I could identify with," "X could be a role model for me," and "X represents someone I would like to be," where "X" is replaced with the role models name and where 0 = "Very Untrue" and 6 = "Very True".

B Role model videos

This figure shows screen shots from the videos used to present the role models (from top left to bottom right: Serena Williams, Nour al Nuaimi, Roger Federer, and Marc Cuban.)



C Tournament entry controlling for beliefs and general factors

This table shows marginal effects of probit models evaluated for a man who chooses the piece rate. All other variables are evaluated at the mean. The tournament-entry decision of Round 3 is the dependent variable (1 for tournament and 0 for piece rate). Panel A (B) shows the coefficients estimated for subjects who saw the male (female) role model. z-statistics based on robust standard errors are reported in parentheses. All variables are defined in detail in Appendix A. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Male role model

	(1)	(2)	(3)
Female subject	–0.171***	–0.149***	–0.138***
	(–2.77)	(–2.86)	(–2.68)
Guessed tournament rank	–0.114***		–0.037
	(–3.19)		(–1.25)
Choice 2		0.656***	0.624***
		(7.28)	(6.42)
Piece rate performance (win)	0.002	–0.006	–0.007
	(0.33)	(–0.89)	(–1.09)
Performance difference 2−1	0.006	0.001	0.001
	(0.36)	(0.08)	(0.11)
Subject education	0.018	0.026*	0.028*
	(1.05)	(1.78)	(1.93)
Importance of math	–0.014	–0.012	–0.013
	(–0.71)	(–0.78)	(–0.83)
Pseudo R^2	0.136	0.339	0.347
Observations	160	160	160

Panel B: Female role model

	(1)	(2)	(3)
Female subject	–0.043	–0.033	–0.031
	(–0.63)	(–0.53)	(–0.51)
Guessed tournament rank	–0.098**		–0.025
	(–2.36)		(–0.63)
Choice 2		0.577***	0.549***
		(6.08)	(5.14)
Piece rate performance (win)	0.007	0.010*	0.010
	(1.02)	(1.76)	(1.56)
Performance difference 2−1	0.021	0.027*	0.027*
	(1.20)	(1.76)	(1.79)
Subject education	0.033*	0.049***	0.050***
	(1.76)	(2.67)	(2.69)
Importance of math	–0.018	–0.026	–0.026
	(–0.91)	(–1.59)	(–1.55)
Pseudo R^2	0.071	0.217	0.220
Observations	145	145	145

Table 1: Summary statistics

Panel A of this table shows summary statistics (mean, standard deviation (sd), median (p50), 1st percentile (p1), 99th percentile (p99), and number of observations (N)) of all subjects in our experiment. Panel B reports means for all female subjects (column (1)), means for all male subjects (column (2)), differences in means between female and male subjects (column (3)), and t-statistics (column (4)), respectively. All variables are defined in detail in the Appendix A.

Panel A: Full sample	mean (1)	p50 (2)	sd (3)	p25 (4)	p75 (5)	N (6)
Subject age	4.67	4.00	2.19	3.00	6.00	835
Subject education	5.17	6.00	1.79	3.00	7.00	829
Importance of math	3.77	4.00	1.75	3.00	5.00	385
Math score 1	11.69	11.00	4.95	9.00	14.00	835
Math score 2	13.07	13.00	5.78	10.00	16.00	835
Math score 3	13.73	13.00	5.91	10.00	16.00	835
Performance difference 2−1	1.38	1.00	2.51	0.00	2.00	835
Time for choice 1	9.80	5.00	21.71	3.00	9.00	835
Time for choice 2	6.45	5.00	6.62	3.00	7.00	835
Choice 1	0.22	0.00	0.41	0.00	0.00	835
Choice 2	0.15	0.00	0.36	0.00	0.00	835
Stereotype threat	1.68	1.00	1.75	0.00	3.00	835

Panel B: Gender differences	mean female (1)	mean male (2)	diff. f-m (3)	t-stat (4)
Subject age	4.941	4.402	0.539	3.608
Subject education	5.109	5.219	-0.109	-0.884
Importance of math	3.709	3.811	-0.102	-0.577
Math score 1	11.323	11.890	-0.568	-1.672
Math score 2	12.631	13.401	-0.771	-1.961
Math score 3	13.249	14.113	-0.864	-2.150
Performance difference 2−1	1.272	1.488	-0.217	-1.262
Time choice 1 (in seconds)	8.654	10.964	-2.310	-1.577
Time choice 2 (in seconds)	7.233	6.559	0.674	0.658
Choice 1 (Tournament=1)	0.165	0.275	-0.110	-3.920
Choice 2 (Tournament=1)	0.108	0.195	-0.087	-3.579
Stereotype threat	1.393	1.948	-0.555	-4.677

Table 2: Tournament entry in neutral condition

Panel A of this table shows the percent of men and women who chose the tournament in column (1) and the piece rate in column (2). Column (3) shows the percentage difference between subjects who entered the tournament and those who did not. Panel B of this table shows the mean math scores of female and male subjects conditional on the compensation schemes. Panel C shows marginal effects of probit models evaluated for a man. All other variables are evaluated at the mean. The tournament-entry decision of Round 3 is the dependent variable (1 for tournament and 0 for piece rate). z-statistics based on robust standard errors are reported in parentheses. All variables are defined in detail in Appendix A. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Tournament entry decision			
	Piece rate (1)	Tournament (2)	Diff. (t-p) (3)
Male subjects (percent)	74.71	25.29	-49.42
Female subjects (percent)	85.5	14.47	-71.06
Difference (f-m)			
Pearson chi2(1)=2.94, Pr=0.087			

Panel B: Tournament entry conditional on performance	(1)	(2)	(3)
	mean(score1) (1)	mean(score2) (2)	mean(s2-s1) (3)
Male subjects			
Piece rate	11.35	12.71	1.33
Tournament	12.83	14.49	1.91
Difference (t-p)	1.48***	1.78***	0.58*
Female subjects			
Piece rate	11.24	12.49	1.24
Tournament	11.68	13.33	1.43
Difference (t-p)	0.44	0.84	0.19

Table 2: Tournament entry in neutral condition cont'd

Panel C: Probit regressions of tournament entry	(1)	(2)	(3)
Female subject	–0.108*	–0.110*	–0.164**
	(–1.75)	(–1.75)	(–2.05)
Piece rate performance (win)		0.007	0.012
		(1.06)	(1.45)
Performance difference 2−1		0.011	0.001
		(0.82)	(0.08)
Subject education			–0.055***
			(–2.58)
Importance of math			–0.005
			(–0.20)
Pseudo R^2	0.018	0.032	0.131
Observations	163	160	81

Table 3: Do role models encourage competitive behavior?

Panel A of this table shows mean ratings for competitiveness and likeability of the role model in columns (1) and (2), respectively, as well as the mean perceived similarity between the the subject and the role model for female participants. The Table also shows the t-statistics for this value to be unequal to 3 (the middle of the scale). Moreover, the table shows the difference between the ratings of the male and female role models. Panel B shows the same values for male subjects. All variables are defined in detail in Appendix A.

Panel A: Women			
	Competitiveness of model (1)	Likeability of role model (2)	Similarity (3)
Male role models	4.600	4.947	2.327
(Different from 3)	20.911	23.519	-5.751
Female role models	4.592	5.151	2.096
(Different from 3)	19.152	29.266	-7.242
Difference (m-f)	0.008	-0.203	0.231
t-stat difference	0.071	-1.835	1.350
Panel B: Men			
	Competitiveness of model (1)	Likeability of role model (2)	Similarity (3)
Male role models	4.857	4.751	2.994
(Different from 3)	29.347	20.300	-0.045
Female role models	4.435	4.968	2.629
(Different from 3)	18.108	29.003	-2.845
Difference (m-f)	0.422	-0.217	0.366
t-stat difference	4.167	-1.970	2.028

Table 4: Tournament entry in role model conditions

This table shows marginal effects of probit models evaluated for a man. All other variables are evaluated at the mean. The tournament-entry decision of Round 3 is the dependent variable (1 for tournament and 0 for piece rate). Panel A shows the coefficients estimated for subjects who saw the male role model and Panel B for subjects who saw the female role model. z-statistics based on robust standard errors are reported in parentheses. All variables are defined in detail in Appendix A. ∗∗∗, ∗∗, and ∗ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Male role model	(1)	(2)	(3)
Female subject	–0.147***	–0.142***	–0.206***
	(–3.39)	(–3.28)	(–3.27)
Piece rate performance (win)		0.010**	0.011
		(2.06)	(1.62)
Performance difference 2−1		0.014	0.002
		(1.35)	(0.15)
Subject education			0.009
			(0.47)
Importance of math			–0.012
			(–0.59)
Pseudo R^2	0.031	0.053	0.084
Observations	343	341	160

Panel B: Female role model	(1)	(2)	(3)
Female subject	–0.075	–0.070	–0.054
	(–1.65)	(–1.53)	(–0.78)
Piece rate performance (win)		0.007	0.011
		(1.27)	(1.54)
Performance difference 2−1		0.009	0.015
		(0.72)	(0.86)
Subject education			0.026
			(1.37)
Importance of math			–0.018
			(–0.93)
Pseudo R^2	0.007	0.013	0.034
Observations	341	337	145

Table 5: Impact of role models on female subjects conditional on performance

This table shows results of probit regressions with tournament entry as the dependent variable. This variable is equal to one if a subject chooses the tournament compensation scheme in Round 3, and zero otherwise. In column (1) of Panel A (B), the sample is restricted to all female subjects in the highest (lowest) performance quartile according to subjects' performance in Round 1. The column presents marginal effects evaluated at a non-female role model (i.e. for the neutral or male role model condition), whereas all other variables are at their mean. In column (2) of Panel A (B), the sample is restricted to all male subjects in the highest (lowest) performance quartile according to subjects' performance in Round 2. The column presents marginal effects evaluated at a non-male role model (i.e. for the neutral or female role model condition), whereas all other variables are at their mean. Results in column (3) of Panel A (B) show coefficient estimates and are based on men and women in the highest (lowest) performance quartile according to subjects' performance in Round 2. All variables are defined in detail in Appendix A. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5: Impact of role models on female subjects conditional on performance cont'd

Panel A: Highest performance quartile	Only female subjects (1)	Only male subjects (2)	All subjects (3)
Female role model	0.168*		–0.353
	(1.83)		(–1.31)
Male role model		0.118	
		(1.23)	
Female subject			–0.658**
			(–2.18)
Female role model X Female subject			0.809*
			(1.87)
Piece rate performance (win)	0.002	0.032***	0.065**
	(0.14)	(2.59)	(1.98)
Performance difference 2−1	0.035	–0.003	0.042
	(1.59)	(–0.15)	(0.96)
Subject education	–0.019	–0.015	–0.049
	(–0.70)	(–0.60)	(–0.84)
Pseudo R^2	0.058	0.063	0.064
Observations	73	104	177

Panel B: Lowest performance quartile	Only female subjects (1)	Only male subjects (2)	All subjects (3)
Female role model	0.013		–0.032
	(0.22)		(–0.12)
Male role model		0.064	
		(0.91)	
Female subject			–0.458*
			(–1.85)
Female role model X Female subject			0.053
			(0.13)
Piece rate performance (win)	–0.033***	–0.007	–0.072*
	(–2.71)	(–0.43)	(–1.69)
Performance difference 2−1	0.008	0.011	0.049
	(0.69)	(0.72)	(1.14)
Subject education	0.023	–0.000	0.034
	(1.33)	(–0.00)	(0.64)
Pseudo R^2	0.113	0.010	0.043
Observations	118	140	258

Table 6: Subjective relative performance predictions and stereotype threat

This table shows coefficient estimates of ordered probit regressions of the guessed piece rate rank on the perceived stereotype threat and a dummy for the subject having seen a female role model. The dependent variable is subjects' guessed performance rank in the piece rate compensation scheme (Round 1) ranging from 1 to 4, where 1 represents the best performer and 4 the worst performer. The sample is restricted to all female subjects. All variables are defined in detail in Appendix A. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Female subjects	(1)	(2)	(3)
Female role model X Stereotype threat			–0.124*
			(–1.74)
Stereotype threat		0.028	0.079*
		(0.81)	(1.77)
Female role model	–0.257*		–0.094
	(–1.65)		(–0.56)
Male role model	–0.192		–0.203
	(–1.27)		(–1.34)
Piece rate performance	–0.085***	–0.084***	–0.083***
	(–5.56)	(–5.55)	(–5.40)
Pseudo R^2	0.041	0.039	0.045
Observations	406	406	406

Table 7: Impact of role models on female subjects conditional on stereotype threat

This table shows results of probit regressions with tournament entry as the dependent variable. This variable is equal to one if a subject chooses the tournament compensation scheme in round three, and zero otherwise. We restrict the sample to only female subjects. Columns (1) and (2) of this table present marginal effects evaluated at a non-female role model and having chosen the piece rate in Round 4. All other variables are evaluated at the mean. Stereotype threat is measured on a scale from zero to 6 using the item: "Men are better at solving math related problems." In column (1), we show results only for subjects that have a stereotype threat score below or equal to the median. In column (2), we show results only for subjects that have a stereotype threat score above the median. Results in column (3) are based on all female subjects. All variables are defined in detail in Appendix A. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 7: Impact of role models on female subjects conditional on stereotype threat cont'd

	Low stereotype threat (1)	High stereotype threat (2)	All (3)
Female role model	0.003	0.109*	–0.053
	(0.06)	(1.80)	(–0.26)
Stereotype threat			–0.141*
			(–1.88)
Female role model X Stereotype threat			0.177*
			(1.70)
Piece rate performance (win)	0.004	–0.002	0.004
	(0.71)	(–0.23)	(0.19)
Performance difference 2−1	0.012	0.011	0.060
	(1.06)	(0.85)	(1.39)
Subject education	0.006	–0.020	–0.022
	(0.49)	(–1.24)	(–0.46)
Guessed tournament rank	–0.020	0.008	–0.029
	(–0.68)	(0.28)	(–0.28)
Choice 2	0.568***	0.331***	1.539***
	(5.79)	(2.65)	(6.85)
Constant			–1.084**
			(–2.43)
Pseudo R^2	0.204	0.132	0.169
Observations	248	152	400

Table 8: Robustness checks

This table shows results of probit regressions with tournament entry as the dependent variable. This variable is equal to one if a subject chooses the tournament compensation scheme in round three, and zero otherwise. Results in column (1) are based on all subjects. In column (2), we show results only for subjects that have a stereotype threat score below or equal to the median. In column (3), we show results only for subjects that have a stereotype threat score above the median. Results in column (3) are based on all female subjects. All variables are defined in detail in Appendix A. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 8: Robustness checks cont'd

	All	Highest Task 2 Performance Quartile	Person is seen as role model
	(1)	(2)	(3)
Female subject X Female role model	0.620**	3.231***	0.837**
	(1.96)	(4.27)	(1.98)
Female role model	–0.303	–1.731***	–0.541*
	(–1.40)	(–3.48)	(–1.81)
Female subject	–0.565***	–0.969*	–0.627**
	(–2.82)	(–1.92)	(–2.49)
Piece rate performance (win)	0.029*	0.055	0.026
	(1.71)	(1.20)	(1.17)
Performance difference (t-p)	–0.001	0.068	–0.045
	(–0.04)	(1.05)	(–1.02)
Subject education	0.010	–0.238***	–0.034
	(0.25)	(–2.70)	(–0.64)
Importance of math	–0.046	–0.142	–0.093
	(–1.03)	(–1.36)	(–1.48)
Guessed tournament rank	–0.296***	–0.644**	–0.323***
	(–3.05)	(–2.33)	(–2.64)
Time choice 1 (in seconds)	0.015**	0.028*	0.017**
	(2.34)	(1.89)	(2.25)
Pseudo R^2	0.096	0.345	0.114
Observations	386	77	246

Figure 1: Tournament entry decision in neutral condition

This figure shows the fraction of female and male subjects sorted in performance quartiles, respectively, entering the tournament in round three without being shown a role model ("neutral" condition). Performance quartiles are computed based on subjects' performance in Round 2 (tournament condition) or based on subjects' performance in Round 3 (subject chooses tournament or piece rate).

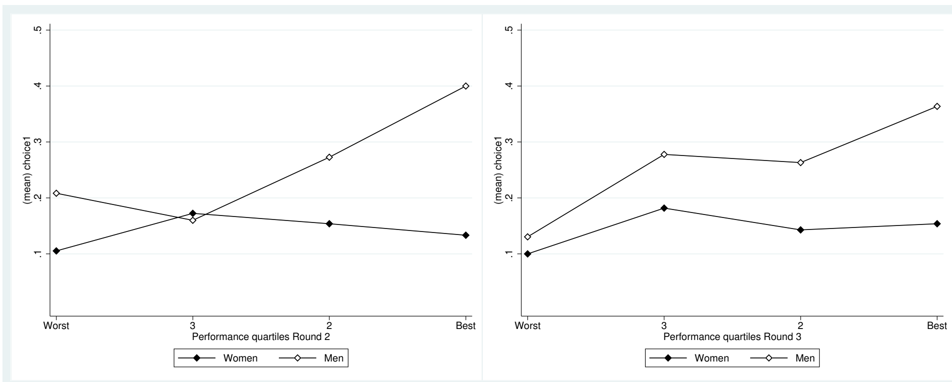
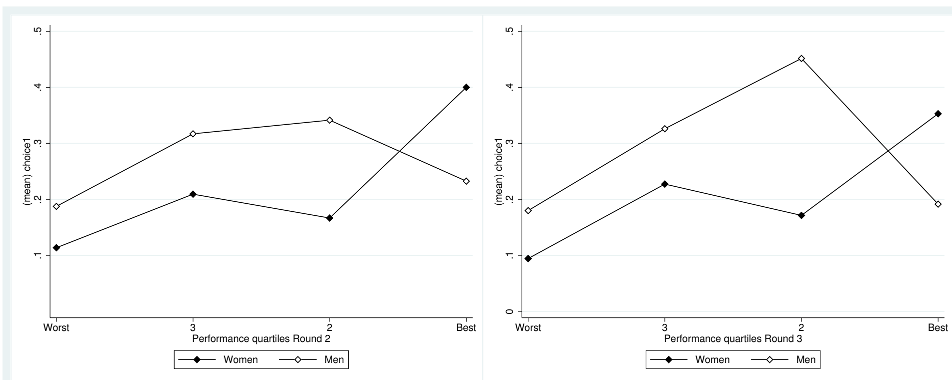


Figure 2: Tournament entry decision with role models

Panel A of this figure shows the fraction of female and male subjects sorted in performance quartiles, respectively, entering the tournament in round three after observing a female role model. Performance quartiles are computed based on subjects' performance in Round 2 (tournament condition) or based on subjects' performance in Round 3 (subject chooses tournament or piece rate). Panel B of this figure provides the same information for subjects who observed a male role model.

Panel A: Female role models



Panel B: Male role models