Hans Böckler Stiftung
Mitbestimmung · Forschung · Stipendien

WORKING PAPER FORSCHUNGSFÖRDERUNG

Nummer 062, March 2018

Gender, Competition and the Effect of Feedback and Task: An Experiment

Alexandra Baier, Brent J. Davis, Tarek Jaber-Lopez and Michael Seidl

FORSCHUNGS-
FÖRDERUNG

The authors are affiliated with the Department of Public Finance, University of Innsbruck, Austria.

We acknowledge the Hans Böckler Foundation for the financial support in running the experiments.

Corresponding author:
Tarek Jaber-Lopez
Department of Public Finance
University of Innsbruck
Universitätstraße 15
A-6020 Innsbruck, AUSTRIA
tarek.jaber-lopez@uibk.ac.at

JEL code: C91, D03
Keywords: competition, gender gap, experiment, feedback

ISSN 2509-2359

Contents

Abstract

While the last decade has shown great advances by women in the labor market, there are still many women who are vastly underrepresented in upper echelon of firms and earn lower salaries on aggregate. The literature suggests gender differences in preferences for competition, confidence, or risk aversion as possible explanations. The present study uses a laboratory experiment to investigate preferences for competition, the effect of feedback and information in a competitive setting when individuals can choose their task. Results indicate that providing ranking feedback and giving participants the option to compete in their preferred task encourages more women to enter competition. Furthermore, feedback affects reported beliefs and improves efficiency. This study provides evidence that manipulating the type of feedback might be an effective alternative to close the gender gap by increasing efficiency in an environment when individuals can choose their preferred task.

1. Introduction

While women have had large gains in education and are better represented in the upper echelon of firms over the last decades, we still observe drawbacks for women in the labor market outcomes. There still exist vast differences between men and women in regards to labor market outcomes, wages, and the number of women in senior management positions (Goldin, 2014). The most common explanation for these gender disparities are differences in competitive labor market preferences leading to differences in labor choices between men and women (Black & Strahan, 2001). A recent overview of these differences in the United States finds little effect from human capital factors like education or experience on the observed gender wage gap (Blau & Kahn, 2017). For example, last year, only 4.8 percent of CEOs of Fortune 500 firms were women.[1] In Europe, the average unadjusted gender wage gap is approximately 16.30 percent.[2]

Besides differences in ability or preferences, the literature suggests women's aversion to compete against men as an explanation for the gender gap in labor market outcomes (Gneezy et al., 2003; Niederle & Vesterlund, 2007). Competitiveness is found to be a crucial predictive factor of labor market outcomes (Kleinjans, 2009). Approaches to address these differences in competitiveness include affirmative action programs such as quotas or preferential treatment for women (Balafoutas & Sutter, 2012; Niederle et al., 2013). However, given the use of positive discrimination of women, affirmative action policies are often difficult to implement. It is suggested that these interventions might interfere with the labor market (Calsamiglia et al., 2013). An alternative approach is altering the information workers receive to induce more women to enter competitive environments (Ertac & Szentes, 2011), as this is not a policy with positive discrimination, it is unlikely to distort the labor market.

The literature suggests that individual preferences, such as risk attitudes, distributional preferences, and differences in confidence as explanations for the gender gap in men entering competitive environments more than women (Balafoutas et al., 2012; Dohmen & Falk, 2011; Kamas & Preston, 2012; Veldhuizen, 2016). Furthermore, there is evidence that stereotype-threats influence women's decision to enter com-

1 Data for the Fortune 500 list is from *Female Fortune 500 CEOs Are Poised to Break This Record in 2017*, available at http://fortune.com/2016/12/22/female-fortune-500-ceos-2017/ (accessed January 21, 2018).

2 www.ec.europa.eu/eurostat/statistics-explained/index.php/Gender_pay_gap_statistics (accessed December 6, 2017).

petition (Dreber et al., 2014; Günther et al., 2010). Activating a certain stereotype can affect the performance of the players. It has a positive effect on the positively stereotyped group and a negative effect on the negatively stereotyped group (Günther et al., 2010; Steele, 1997).

This paper makes two important contributions. First, we let subjects compete in their preferred task. We believe that this setting represents a more realistic environment that matches the labor market, as entry into the labor market is endogenous. Furthermore, this addresses the question on whether stereotype-threats affect subject's decisions to enter into competition. In previous experiments, subjects perform in only one task, or they perform in two different tasks but do not have the option to choose in which one they prefer to compete (Große & Riener, 2010; Shurchkov, 2012; Wozniak et al., 2014). In our experimental environment subjects have the option to endogenously choose between two tasks, a stereotypically-male task and a stereotypically-female task (Günther et al., 2010). This mechanism enables us to analyze two different types of subjects, those who perform in their preferred task and those who are assigned to their non-preferred task (the details are described in Section 3). This is often the case in real-life, individuals are reluctant to apply for non-desired tasks or jobs and therefore they may exert lower effort levels. An illustrative example is an unemployed person who has to fulfil certain requirements to receive their benefits or in economic duress where an individual needs to accept any job in order to avoid unemployment. This creates an interesting environment where we can study if subjects perform better (worse) in their preferred (non-preferred). Second, we compare a control treatment in which subjects receive information on their own performance to two additional treatments in which we vary the type of feedback. The first type of feedback gives a detailed information table on the performance of all subjects within the same group. This information allows a perfect self-assessment on the relative performance level among the competitors and the ranking. The second type of feedback contains additional information on the gender of the competitors in addition to their corresponding performances. Therefore, we investigate whether or not providing feedback on only relative performance and/or additional information on the subjects' gender reduces the gender gap.

Thus, our contribution to the experimental literature on gender and competition is threefold:

We use a novel setting in which we let subjects self-select in their favorite task and give a better view on how a feedback mechanism as an alternative to affirmative action policies.

We analyze two factors that influence competitive entry rates that are likely to be affected by our treatment manipulations: confidence and efficiency.

Finally, we perform an individual level analysis to seek for personal factors that drive an individual's preference for entering competition, such as risk aversion, social preferences and other sociodemographic variables.

The rest of the paper is structured as follows: Section 2 gives a detailed overview of the existing literature in this research field. Section 3 describes the experimental design and procedures. In Section 4 we present the results. Section 5 contains an econometric analysis and Section 6 concludes.

2. Related Literature

2.1 Competitiveness

We were motivated by a wealth of literature that deals with gender differences in competitive environments. Experiments in this field show that, contrary to men, women tend to shy away from competition despite no difference in performance (Balafoutas & Sutter, 2012; Gneezy et al., 2003; Niederle & Vesterlund, 2007). Most of the evidence is based on experiments that use the setup from Niederle and Vesterlund (2007). In terms of performance, there are no significant differences between men and women. Still, when having the opportunity to choose a payment scheme, men choose the tournament twice as often as women. As this results cannot be explained by performance effects, the authors refer to differences in competitiveness between men and women (Niederle & Vesterlund, 2007).

Further experiments replicate the results that men and women perform equally well in all stages but women do not enter the competition as often as men do (Balafoutas & Sutter, 2012; Ertac & Szentes, 2011). The same finding can also be shown in different experimental setups or with different subject groups. For instance, Gneezy et al. (2003) provide evidence from a laboratory experiment where the participants had to solve mazes. Their results show that, relative to women, performance of men significantly increases under competition. Gneezy and Rustichini (2004) conduct a field experiment with 9-10 year old children to examine differences in competitive behavior between boys and girls. During a physical class, the children complete two 40 meter runs on a track. First, they run the track alone and then, paired with another child of similar speed, against each other. While the authors cannot find any difference between boys and girls in the first round, boys improve significantly more than girls in their performance in the competitive environment. These experiments hint that the difference in the competitiveness between men and women is responsible for the observed gender gap in the labor market outcomes.

2.2 Task

Literature suggests that the task itself influences men and women's decision to enter the tournament. Günther et al. (2010) show that a stereotypically male task encourages men, but not women, to perform better.

This pattern cannot be found when using a female or neutral task. Analyzing the entry decision of men and women in a setting that includes a quantitative task (arithmetic) and a verbal task, verifies a high gender association of the tasks. While there is a significant gender gap in the entry rate for the math task this gap is not significant for the verbal task. This result is further replicated in the field when subjects have to perform a male task (ball throwing) and a verbal task (Große & Riener, 2010). An experiment among adolescents in Sweden further confirms that boys are equally likely to compete in a math and word task but the girls shy away from competition in the math task. Thus, the authors find a gender gap in the entry decision in the math, but not in the word task (Dreber et al., 2014).

2.3 Feedback

As a possible way to manipulate the behavior of men and women, many experiments implement feedback mechanisms (Cason et al., 2010; Ertac & Szentes, 2011; Wozniak, 2012). These experiments have differing results when analyzing feedback as an instrument to encourage women to more competitiveness. While performance information among the competitors can help to reduce the gender gap in the entry rate (Ertac & Szentes, 2011; Ewers, 2012), this result cannot be shown for aggregated information (Jeworrek, 2016; Wozniak et al., 2014).

Ewers (2012) shows that after changing the information provided from an aggregated distribution of performances of the opponent, the entry rate decreases significantly and the gender gap vanishes. This result relies on a two-player game in which subjects work on a multiple-choice quiz without time pressure. For a concentration task, Jeworrek (2016) demonstrates that giving aggregated information of the average performance of previous players does not substantially reduce the gender gap in the tournament entry decision. Furthermore, men have a stronger reaction to feedback compared to women. Information that there is no gender gap in ability, might additionally put pressure on the women (Jeworrek, 2016). An experimental approach that implements feedback in the standard design from Niederle and Vesterlund (2007) shows that providing information about the maximum performance in the group from the previous performance lets the gender gap disappear (Ertac & Szentes, 2011).

Contrary to this result, Wozniak et al. (2014) cannot find a feedback effect on the entry rate of women. Providing information about the performance distribution of the whole session and varying the group size

(groups of 2, 4 or 6) reduces the entry share of men but does not affect the entry rate of women. The findings suggest that high ability women still enter competition at a suboptimal low rate. As this result is based on a within-subject experimental design that differentiates two tasks, a math task and a word task, it can be concluded that feedback barely affects women even in the word task, but men substantially change their behavior. This decrease in the men's entry rate reduces the gender gap (Wozniak et al., 2014). Motivated by the mixed results that feedback and information have on competitive entry choices, our experimental setup compares an environment where each subject completes two different type of real effort tasks and two types of feedback in a between subject framework.

2.4 Confidence and Risk

Other individual preferences might influence the behavior in competitive environments as well, such as risk aversion and overconfidence (Niederle & Vesterlund, 2010; Veldhuizen, 2016). The risk attitudes literature shows that women are more risk-averse than men (Croson & Gneezy, 2009; Eckel & Grossman, 2008). These risk preferences only partly explain the obtained gender gap in competition entry (Muriel Niederle & Vesterlund, 2007). In a similar setup, but with lottery choices instead of payment choices, van Veldhuizen (2016) shows that risk aversion rather than competitiveness is the major driving factor in competition entry decisions. Secondly, van Veldhuizen (2016) finds that differences in overconfidence between men and women is a highly important factor in competitive entry decisions. Kamas and Preston (2012) prove that men are more overconfident than women and that this difference explains the largest part of the gender gap in competition entry. Distinguishing three types of confidence levels, they show that stating the placement relative to the other competitors, thus giving expected rank, is the measure that predicts the confidence level best. When controlling for overconfidence, no gender gap is found in the tournament entry decision (Kamas & Preston, 2012).

We contribute to the existing literature by first addressing the aforementioned stereotype-threat. To the best of our knowledge, we are the first to let subjects choose their preferred task and consequently whether to compete on it. Implementing a choice of task between a female and a male task will provide more insights on whether there is indeed a stereotype-threat that affects the entry decision of men and women. Second, we use an economic laboratory experiment to investigate the effect of

providing information on subjects' ranking in a competition and others' performance. Additionally, we provide information on the gender of each competitor.

3. Experimental Design

The experiment was conducted in the EconLab at the University of Innsbruck in October and November 2017 with the software Z-tree (Fischbacher, 2007). The subject pool consisted of 360 students (180 men and 180 women). Subjects were recruited via HROOT (Bock et al., 2014). The sessions took approximately 60 minutes and the average payoff was 14 Euro. We recruited men and women separately to ensure perfectly balanced groups by gender. We also only recruited subjects who were native German speakers due to the presence of the word task.

3.1 Stage 1: Piece Rate

Each subject participates first in the piece rate math task and subsequently in the piece rate word task. The math task consisted in adding as many sets of five two-digit numbers as possible within 3 minutes Niederle & Vesterlund, 2007). Subjects receive 0.50 Euro for each correctly solved calculation. The word task consisted in an own-developed letter puzzle. Subjects have to build German verbs in the infinitive tense from a set of six letters in random order. Subjects have three minutes to solve as many words as possible. Subjects receive 0.50 Euro for each correctly written word.

3.2 Stage 2: Tournament

Each subject participated first in the tournament math task and subsequently in the tournament word task. Each tournament is composed of three parts: the task, the belief elicitation and the feedback which varies depending on the treatment. In the tournament math task, subjects compete against each other in the math task in a group of six. The two best performers were paid 1.50 Euro for each correctly solved calculation while the other four competitors received nothing with ties broken randomly. Thereafter, subjects state then their guessed rank in the group (between 1 and 6), which was announced prior to the tournament. These subjective guesses are a valid measure for confidence levels of the subjects (Kamas & Preston, 2012). Correct guesses were rewarded with 0.5 Euro each. Subjects were only informed about the outcome of their guesses at the end of the experiment. Subjects receive feedback information on performance in the tournament math task according to the

treatment (details in section 3.7). In the tournament word task, subjects compete against each other in the word task in a group of six. The two best performers were paid 1.50 Euro for each correctly solved word while the other four competitors get nothing. Subjects state then their guessed rank in the group (between 1 and 6). Correct guesses were rewarded with 0.5 Euro each. Subjects were only informed about the outcome of their guesses at the end of the experiment. Subjects receive feedback information on performance in the tournament word task according to the treatment (details in section 3.7).

3.3 Stage 3: Choice of Task

Participants state priorities on which task they prefer for the upcoming rounds. These priorities rank from “Math strongly preferred”, “Math weakly preferred”, Indifferent”, “Word weakly preferred” to “Word strongly preferred”. According to these priorities, subjects are assigned to the most preferred task as long as the gender-balanced group composition is ensured. Otherwise, they are assigned to the non-preferred task, ties are broken randomly.

3.4 Stage 4: Forced Tournament

Subjects are forced to play a tournament. They perform either in the chosen task or in the task they got randomly assigned. Again, the two best players win and get 1.50 Euro for each correct answer while the rest gets nothing. Subjects are then asked about their beliefs (which were again unannounced) and they receive feedback information on performance in the forced tournament according to the treatment (details in section 3.7).

3.5 Stage 5: Choice of payment scheme

Subjects are asked if they want to perform in their chosen/assigned task under piece rate scheme or tournament. If they choose tournament their performance is compared to the performance of the other group members from the previous Stage (*Stage 4: Forced Tournament*). In every treatment subjects only receive information on their own performance in this stage. Subjects receive 0.50 Euro for each correctly written word.

3.6 Stage 6: Individual preferences and sociodemographic questionnaire

In this last stage we ask subjects about their attitudes towards risk by using the bomb risk elicitation (Crosetto & Filippin, 2013). We also obtain their distributional preferences via the equality equivalence test (Kerschbamer, 2015). In a final sociodemographic questionnaire, subjects are asked about their age, gender, field of study, their education path, and the education of their parents.

3.7 Treatments

We have three different treatments, which are described as follows:

Treatment 1: Own performance Feedback
In this treatment, subjects receive only information on their own performance, i.e. the number of correct answers in the forgone tournament.

Treatment 2: Ranking Feedback
Subjects receive information on each group member's performance, including their own performance, i.e. number of correct answers and the rank within the group in the forgone tournament. Notice that in this treatment subjects are also aware of their relative performance, i.e. how good or bad they performed compared to the others.

Treatment 3: Ranking Feedback and Gender Information
Subjects receive information on each group member's performance and gender, including their own performance, i.e. number of correct answers, and the rank within the group in the forgone tournament. Notice that in this treatment subjects are also aware of their relative performance, i.e. how good or bad they performed compared to the others. The information of the members' gender led them to know also the performance of those belonging to the same and to the other gender group.

4. Results

The experiment was conducted with a standard subject pool of students. The mean age is 23 (st. dev. 3.81), approximately two thirds of the participants are Bachelor students and one third are Master students. 47.78 percent of the subjects are enrolled in economics and business administration, 35.83 percent in natural sciences and 16.39 percent in humanities.

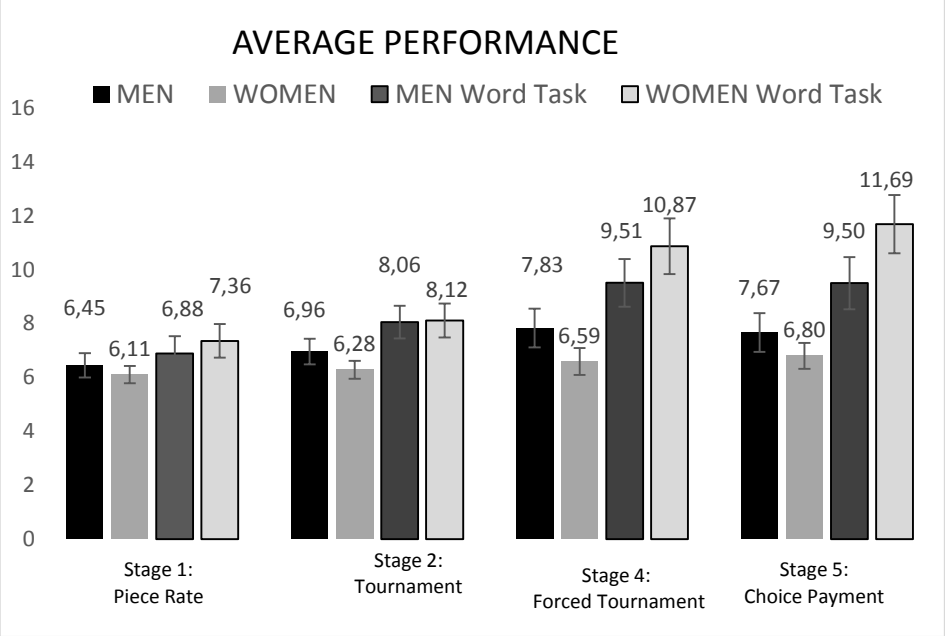
4.1 Performance

Figure 1 displays the average performance levels of men and women in each stage and task. Overall, we do not find significant gender differences of performance for both task in any stage before choosing their preferred task, i.e. in Stage 1 and Stage 2. This result is in line with previous studies such as Niederle & Vesterlund (2007). Some significant results are found in the subsequent stages. In Stage 4, after choice of task, women perform slightly significantly better than men in the word task ($p=0.054$)[3] and men significantly better than women in the math task ($p=0.046$). In Stage 5, when subjects choose the task and the payment scheme, we do not find significant gender differences of performance in the math task on the one hand ($p=0.337$). On the other hand, women perform significantly better than men in the word task ($p=0.003$). A possible explanation for the gender differences in Stage 4 and 5 is that women increase their motivation after choosing their preferred task while men do not.

Furthermore, subjects solved significantly more word problems than math problems overall. On average the performance in the word task is significantly higher than in the math Task in the piece rate stage ($p=0.002$). This result is confirmed in the tournament stage ($p<0.001$), in the forced tournament ($p<0.001$), and in the choice of the payment scheme stage ($p<0.001$). Despite these differences, many subjects choose the math task.

3 Henceforth, p-values refer to Mann–Whitney U test when analyzing numerical data and Chi-Square (χ^2) test when using categorical data.

Figure 1: Average performance in each stage, by gender. Bars represent the average number of correct answers in each task and in each stage.



Furthermore, neither women nor men improve their performance when comparing Stages 1 and 2 in the math task (p=0.103) or between Stages 2 and 4 (p=0.154), or between Stages 4 and 5 (p=0.122). The same pattern can be shown for the learning effect in the word task (p=0.396; 0.365; and 0.161 respectively). Overall, it can be concluded that there is no substantial gender gap, neither in the performance in the tasks nor in any improved performance effects over the rounds.

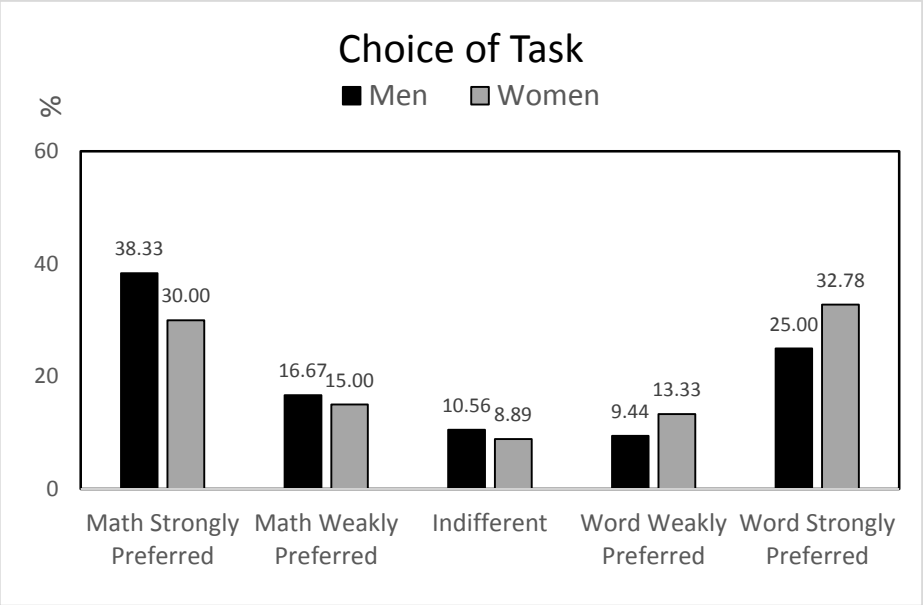
An interesting question refers to the implementation of feedback and the changes of upcoming performance. We do not find significant differences between men and women in performance levels in the math and the word task across the different treatments (p=0.15 and p=0.60 in second treatment; p=0.27 and p=0.08 in third treatment for math and word performance in the chosen task stage, respectively). Thus, men and women react in an equal manner to relative feedback in terms of performance. The choice of task allows us to differentiate between subjects that got their preferred task and those that were assigned their non-preferred task. Analyzing the performance of these two groups gives the opportunity to test whether subjects might get motivated or demotivated irrespectively after (not) getting assigned to the preferred task. In the first performance stage after the choice of task (Stage 4) we can observe significant differences between subjects that got their preferred task and those that did not (p=0.04 for the math task; p=0.001 for the word task).

This performance difference likely arises from the fact that subjects that did not get their preferred task had to perform in their weaker task. The improvement in the performance after the choice of task does not significantly differ between the subjects that were assigned to the preferred task and the subjects that were assigned to their non-preferred task in Stage 4 (p=0.629 for math and p=0.977 for word). Consequently, there is no evidence for effects like demotivation or disappointment. Subjects that stated indifference (N=35) are omitted from the analysis. It is notable that most of the subjects got their preferred task (N=280) and thus the result above relies on a small sample (N=45).

4.2 Choice of Task

Figure 2 displays the stated priorities of men and women for each task. Men prefer the math task significantly more than women (p=0.058), and women have significantly higher preferences for the word task than men (p=0.024). There are some significant differences between treatments for men. Men choose the word task significantly more in treatment 2 than in treatment 1 (45 percent vs. 28 percent; p=0.058). They choose slightly significantly less the word task in treatment 3 than in treatment 2 (30 percent vs. 45 percent, p=0.090). Furthermore, they choose significantly more often the math task in treatment 3 than in treatment 2 (67 percent vs. 45 percent, p=0.017). Therefore, the type of feedback received in each treatment seems to have an effect on men. When they just receive feedback on ranking (treatment 2) they seem to prefer more the word task, however, when they receive information on ranking and gender (treatment 3) they have clear preferences for the math task. No significant differences are found for women across treatments.

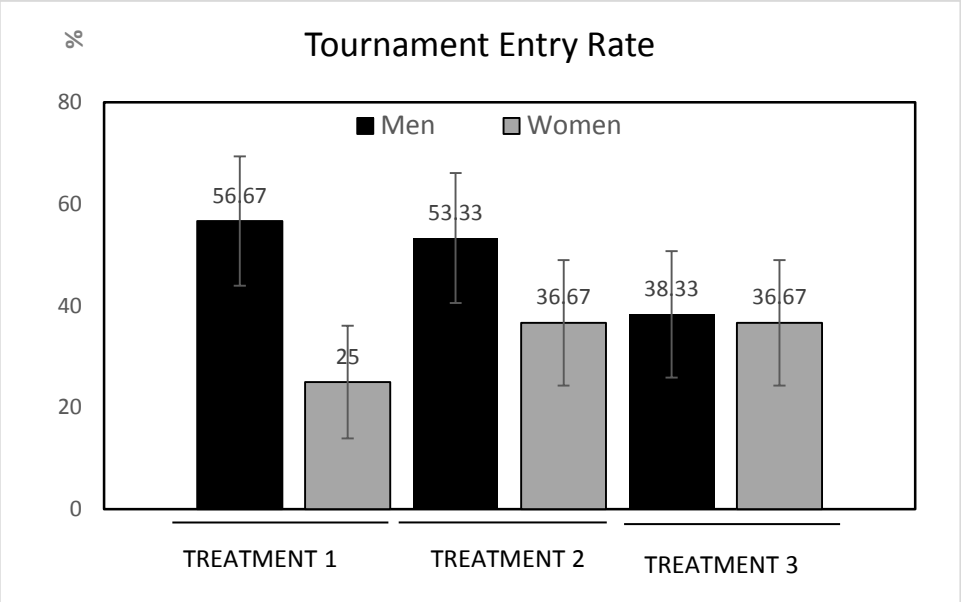
Figure 2: Distribution of preferences for the math and the word task, by gender. Bars represent the percentage of subjects who had the corresponding preferences for each task.



4.3 Entry Rate

Figure 3 shows the tournament entry rate across treatments, in other words the percentage of people who chose to compete as a payment scheme in Stage 5. In treatment 1, 56.57 percent of men and 25 percent of women choose to enter the tournament. This gender difference is highly significant ($p<0.001$). This gap is consistent with the existing literature (Niederle & Vesterlund, 2007). In treatment 2, the share of men entering into competition decreases only slightly to 53.33 percent but the share of women entering competition increases to 36.67 percent. The gender gap becomes lower and the difference between men and women is only significant at the 10 percent level ($p=0.067$). In treatment 3, the share of men entering competition shrinks substantially to 38.33 percent while the share of women entering competition remains constant when compared to treatment 2 at 36.67 percent. In treatment 3, there is no significant difference, and the gender gap is almost non-existent ($p=0.851$).

Figure 3: Tournament entry rate, by gender. Bars represent the percentage of subjects choosing tournament as payment scheme in Stage 5.



Our results confirm findings from earlier experiments demonstrating that feedback can help to reduce the gender gap (Ertac & Szentes, 2011). Contrary to the results from Wozniak et al. (2014), our results indicate that not only do men react to the implemented feedback, but women change their behavior as well. As a robustness check, we do not find significant difference in the entry rate between subjects who were allocated in their preferred task and those who did not (p=0.644). In the next sections, we investigate two factors that might change across treatments and could help explain the smaller of the gender gap in treatments 2 and 3, confidence and efficiency.

4.4 Overconfidence

Literature suggests confidence levels as an important factor in the decision on whether to enter in a tournament (Kamas & Preston, 2012). It is worth investigating if relative feedback and gender information have an effect on the confidence levels of men and women. This section analyzes measures of (over)confidence. The estimated rankings represent the beliefs of the participants on the relative performance in the group. Based on Kamas and Preston (2012) we define as overconfident a sub-

ject that expected to win the tournament. In other words, those subjects who ranked themselves as first or second in the belief question in the forced tournament, Stage 4.

Figure 4: Overconfidence level across treatments, by gender. Bars represent the share of subjects who guessed they will win, i.e. they are ranked in the first or second position.

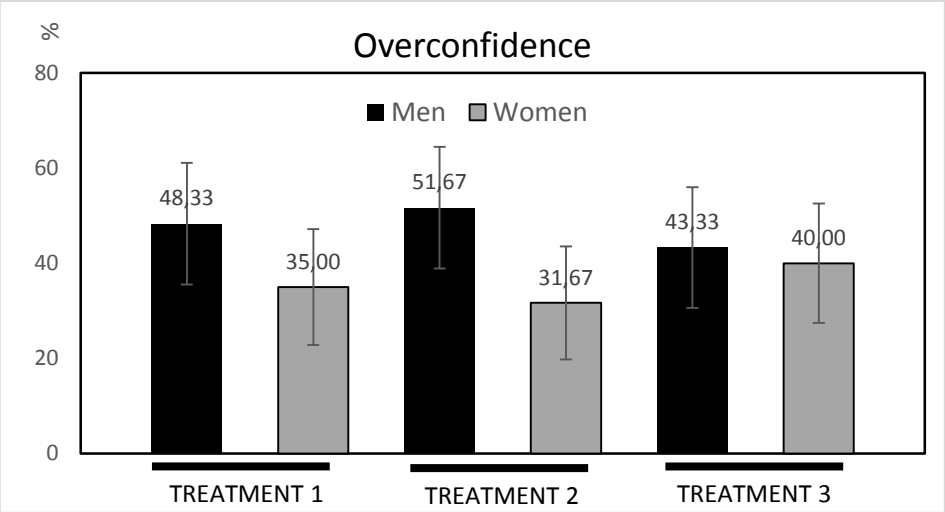


Figure 4 shows the percentage of overconfident men and women across treatments. Overall men are significantly more overconfident than women (71.66 percent vs. 53.33 percent, p=0.005). Men are more overconfident than women in treatment 1, although the difference is not significant (p=0.195). In treatment 2 the differences are significant (p=0.041) and become again insignificant in treatment 3 (p=0.853). Thus, relative feedback with gender information affects both men and women, in their guessed ranks and (over)confidence levels. The aforementioned results suggest that the shrinkage of the gender gap in treatment 3 might be driven by both a decrease of overconfidence among men and an increase of overconfidence among women.

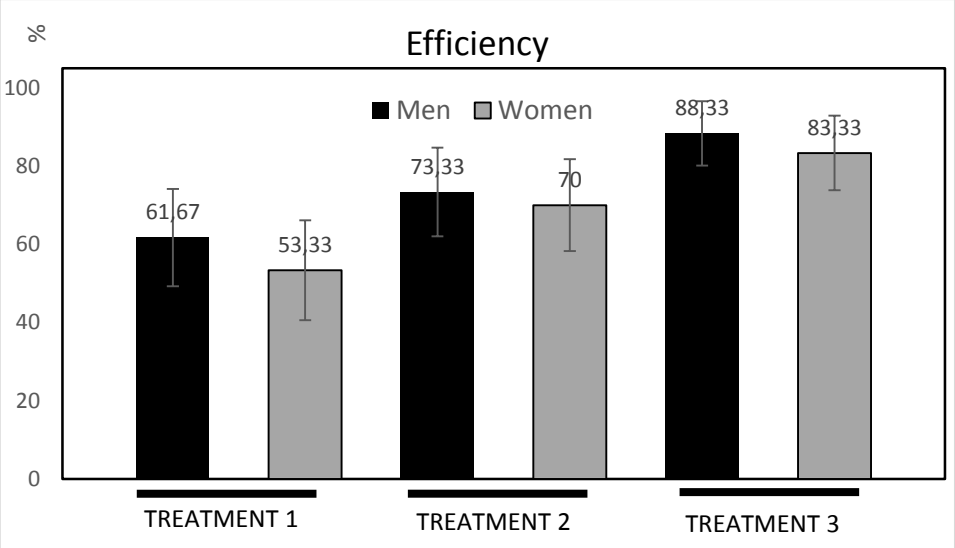
4.5 Efficiency

We classify as efficient those decisions in which subjects maximize their earnings by choosing the adequate payment scheme based on their ability. Therefore, a decision is efficient when a subject chose the tour-

nament and won the tournament or the subject chose piece rate and would not win the tournament. Overall, 74.44 percent of the men and 68.89 percent of the women have made their decision efficient (p=0.351). We do not find significant differences between men and women over treatments (treatment 1: p=0.460, treatment 2: p=0.999, and treatment 3: p=0.601).

Figure 5 shows an increase in the share of efficient decisions across treatments for both men and women. 57.50 percent of subjects make efficient decisions in treatment 1, 71.67 percent in treatment 2, and 85.83 percent in treatment 3. All the three between treatment differences are significant (p<0.04, for all three comparisons). The aforementioned results show that subjects take more efficient decisions when they are aware of their rank and even more when they are aware of their rank and the gender. This result might explain the reduction of the gender gap in treatment 2 and 3. A possible explanation is that in treatment 1 men over-compete and women under-compete but due to the relative ranking men and women are adjusting their abilities to their decision on whether to compete.

Figure 5: Efficient decisions across treatments, by gender. Bars represent the percentage of subjects who took an "efficient" decision in the choice of payment scheme – Stage 5.



4.6 Risk

Previous literature suggests differences in risk attitudes between men and women to account for large parts of the observed gender gap in the entry rate (Niederle & Vesterlund, 2007). Using the bomb risk elicitation task (Crosetto & Filippin, 2013) we find that in general men and women do not differ significantly in their risk attitudes. Regarding the men 56% are risk-averse, 16% risk-neutral and 28% risk-seeking. The women reveal shares of 61%, 19% and 19%, respectively. There is only a slightly significant difference in risk-seeking preferences between men and women ($p=0.048$) and no difference in risk-averse preferences between men and women ($p=0.336$). This result is consistent for all treatments and is therefore not affected by the implementation of relative feedback on forgone performances. However, we can confirm earlier evidence that risk-loving subjects are more willing to enter into competition ($p=0.004$) and risk-averse subjects are less likely to enter ($p=0.006$). Furthermore, the results demonstrate that risk-loving men enter more into competition more than risk-loving women ($p=0.070$) and risk-averse women enter significantly less often in competition than risk-averse men ($p=0.014$).

5. Econometric Analysis: Determinants of the willingness to compete

5.1 Individual preferences

In this section, we analyze if subject's individual preferences such as overconfidence, risk attitudes, and social preferences are potential determinants of their tournament entry decision. Therefore, our dependent variable is a dummy variable, which takes value 1 if the subject decided to enter in tournament in Stage 5, and 0 otherwise. All specifications shown in Table 1 estimates Probit models with the dependent variable *Enter Tournament*. Our control variables are *Female* (1 if women, 0 otherwise). We include dummy variables for *treatment 2* and *treatment 3*, leaving treatment 1 as benchmark. *Guess Win Tournament* and *Guess Win Forced Tournament* represent dummy variables for overconfidence (as defined in section 4.4) in Stage 2 (Tournament) and Stage 4 (Forced Tournament), respectively. *Risk Averse* takes value 1 if the person is risk averse, 0 otherwise. Finally, we control for all the dummy variables measuring social preferences type for both the in-group, i.e. a person from same gender, and the out-group, i.e. a person from the other gender. The social preferences types included are *Selfish*, *Inequity Averse*, *Altruistic* and *Spiteful*, leaving *Inequity Loving* as benchmark (Kerschbamer, 2015).

In column (1) we include only the treatment variables, in column (2) we include overconfidence measures. Risk attitudes are added in column (3), Social Preferences in column (4), and column (5) is our full specification in which we include all of the aforementioned variables. Column (1) shows a large and highly significant gender gap in the willingness to compete; women are on average 43.3 percent less likely to enter in tournament than men are. Interestingly, in column (2) Female becomes almost insignificant and the coefficient decreases, meaning that the gender is roughly robust when controlling overconfidence. Women are 29.5 percent less likely to enter in competition than men, although this is only significant at a confidence level of 10 percent. Additionally, subject's expectation to have won in Stage 2 (*Guess Win Tournament*) and Stage 4 (*Guess Win Forced Tournament*) yields a higher likelihood of entering competition as expected.

In specification (3), we control for risk aversion and the gender gap remains to the same level than in (1) and remains highly significant. This means that the gender gap is robust to including risk. The variable *Risk Averse* is negative and highly significant, thus risk averse people are

less willing to compete. In column (4), we control for the different social preference types, using Inequity Loving as benchmark. Again, we find that women are 42.6 percent less willing to compete than men and this result is highly significant. In our full specification in column (5), we include all controls finding that the gender gap is no longer significant. Therefore, the gender gap in our environment is not robust to including overconfidence, risk preferences, and social preferences. Only providing task choice does not appear to reduce the gender gap, however the combination of task choice and feedback (in treatments 2 and 3) does appear to reduce the gender gap.

The results from these regressions suggest that one of the main reasons behind the gender gap is confidence. Based in our results from section 4.4 we can conclude that a way to tackle the gender gap is through affecting the levels of confidence of both genders. Our treatment manipulation seems to be an effective mechanism in order to adjust individual's confidence level that more actually reflect their abilities.

Table 1: Econometric analysis of individual preferences

	(1)	(2)	(3)	(4)	(5)
	Enter Tournament	Enter Tournament	Enter Tournament	Enter Tournament	Enter Tournament
Female	-0.433***	-0.295*	-0.421***	-0.426***	-0.250
	(0.158)	(0.162)	(0.153)	(0.165)	(0.167)
Treatment 2	0.111	0.184	0.092	0.078	0.164
	(0.117)	(0.146)	(0.126)	(0.128)	(0.162)
Treatment 3	-0.082	-0.083	-0.077	-0.114	-0.089
	(0.130)	(0.156)	(0.140)	(0.131)	(0.165)
Guess Win Tournament		0.718***			0.732***
		(0.177)			(0.191)
Guess Win Forced Tournament		0.992***			1.018***
		(0.143)			(0.150)
Risk Averse			-0.353***		-0.347***
			(0.118)		(0.119)
Selfish in-group				-0.254	0.0658
				(0.235)	(0.280)

Selfish out-group				-0.052	-0.490*
				(0.237)	(0.291)
Inequity Averse in-group				0.784	0.371
				(0.551)	(0.643)
Inequity Averse out-group				-1.147**	-0.874
				(0.463)	(0.538)
Altruistic in-group				0.324	0.028
				(0.343)	(0.334)
Altruistic out-group				-0.629*	-0.433
				(0.353)	(0.387)
Spiteful in-group				0.144	0.239
				(0.447)	(0.464)
Spiteful out-group				-0.611	-1.048*
				(0.574)	(0.583)
N	360	360	360	360	360

*Marginal effects; Standard errors in parentheses (d) for discrete change of dummy variable from 0 to 1. * p < 0.10, ** p < 0.05, *** p < 0.01*

5.2 Sociodemographic Background

In this section, we analyze the effect of subjects' sociodemographic background on the willingness to enter into tournament. In addition to individual preference, vast variability heterogeneity in family backgrounds may have a large impact on preferences for competition. As most experimental studies do not include this information, we gather sociodemographic and data from a post-experiment questionnaire. In the regression analysis, the dependent variable is a dummy variable, which takes value 1 if the subject decided to enter in tournament in Stage 5, and 0 otherwise. All specifications shown in Table 2 estimate Probit models with the dependent variable *Enter Tournament*. Our control variables are extracted from the questionnaire at the end of the experiment. For presentation purposes, we classify the variables in seven blocks from which there is always a benchmark variable as reference and the rest remain the dummy variables. In Block I, we control the household conditions, and we use *Both parents live together* as benchmark, with

additional categories of *living with only their mother* (or *father*), and *other*. In Block II, we control the parent's working conditions and we use *Only father works full time* as benchmark, with additional categories of only the *mother works full time, both parents work fulltime*, *part-time father and fulltime mother*, and *other*. In Block III, we display Mothers' education with *School* as benchmark, with additional categories of *Apprenticeship*, completed *High School (academic track), Bachelor degree*, *Master's degree*, and *Ph.D*. Similarly, in Block IV we control Father's education with *School* as benchmark with additional categories being the same as in Block III. In Block V and VI we control for the type of job of the mother and father, respectively, using *Public Sector* as benchmark. Additional job categories in Block V and VI are *self-employed, academic*, and *service*. In Block VII, we control for subjects' political orientation by using *Left Wing* as benchmark, with additional categories of *Right Wing*, *Moderate*, and *Somewhat Right or Left Wing*. In the specification (1) we perform our regression for the whole sample, in column (2) we include only men and in column (3) we include only women.

We find a positive and significant effect in Block II suggesting that Parents' working status matters. In column (3) we observe that in a household where both parents work full time, women are 59.4 percent more likely to compete than in a household where only the father works full time. In Block III, the variable *Ph.D.* has a significant and negative effect for men (column 2) and positive for women (column 3). It seems that mother's education has a negative impact on men and positive impact on women. Men whose mother has a Ph.D. are 125.2 percent less willing to compete than men whose mother has only a school level of education. Contrary, women whose mother has a PhD are 132.6 percent more willing to compete than women whose mother has only school level of education.

Our results suggest that the sociodemographic background has a certain impact on women's willingness to compete. Those women who have as reference a model in which both parents are working are more willing to compete. Furthermore, mothers' education has also a positive impact on women, suggesting that women whose mother has a high level of education are more willing to compete as well. We find no effect that individual political attitudes have on preferences for competition.

Table 2: Econometric analysis of sociodemographic background

	(1) All	(2) Men	(3) Women
	Enter Tournament	Enter Tournament	Enter Tournament
Block I-Household conditions Benchmark: Both parents live together			
Only mother	-0.139	0.130	-0.184
	(0.205)	(0.342)	(0.300)
Only father	-0.450	-0.219	-0.679
	(0.433)	(0.652)	(0.739)
Other	-0.119	0.0130	-0.197
	(0.188)	(0.272)	(0.318)
Block II-Parents' working status Benchmark: Only father works full time			
Fulltime mother	-0.0482	-0.620	0.217
	(0.390)	(0.641)	(0.572)
Full time both parents	0.236	-0.235	0.594**
	(0.164)	(0.249)	(0.257)
Part time father full time mother	-0.501	0	0
	(0.716)	(.)	(.)
Other	0.380	0.0967	0.521
	(0.238)	(0.373)	(0.380)
Block III-Mother's education Benchmark: School			
Apprenticeship	-0.252	-0.223	-0.0501
	(0.245)	(0.378)	(0.377)
High school (academic track)	-0.307	-0.328	-0.0123
	(0.281)	(0.417)	(0.466)
Bachelor	-0.0701	-0.488	-0.160
	(0.391)	(0.578)	(0.698)
Master	-0.373	-0.135	-0.258
	(0.297)	(0.487)	(0.460)

Ph.D.	-0.250	-1.252**	1.326*
	(0.390)	(0.566)	(0.759)
Other	-0.0802	0.256	0.0475
	(0.481)	(0.941)	(0.676)
Block IV-Father's Education Benchmark: School			
Apprenticeship	-0.0937	-0.664	0.178
	(0.275)	(0.454)	(0.417)
High school (academic track)	0.0897	-0.490	0.343
	(0.316)	(0.517)	(0.474)
Bachelor	-0.142	-0.392	-0.251
	(0.427)	(0.595)	(0.824)
Master	0.00230	-0.651	0.251
	(0.312)	(0.504)	(0.471)
Ph.D.	0.311	0.223	0.723
	(0.362)	(0.576)	(0.553)
Other	-0.468	-0.517	-0.418
	(0.410)	(0.610)	(0.639)
Block V-Type of job mother Benchmark: Public Sector			
Self employed	0.0548	0.310	-0.0402
	(0.210)	(0.316)	(0.332)
Academic	0.297	0.694	0.163
	(0.312)	(0.488)	(0.523)
Services	-0.124	-0.255	-0.117
	(0.204)	(0.330)	(0.296)
Other	0.170	0.0795	0.123
	(0.212)	(0.325)	(0.320)
Block VI-Type of job father Benchmark: Public Sector			
Self employed	-0.197	-0.618**	0.0300
	(0.203)	(0.314)	(0.317)
Academic	-0.317	-0.478	-0.647
	(0.259)	(0.378)	(0.461)
Service	-0.181	-0.362	-0.104
	(0.207)	(0.321)	(0.315)
Others	-0.738***	-1.099***	-0.646
	(0.286)	(0.416)	(0.454)

Block VII-Political orientation Benchmark: Left Wing			
Right Wing	-0.303	0	0.390
	(0.980)	(.)	(1.093)
Rather Right Wing	-0.215	-0.261	0.208
	(0.321)	(0.446)	(0.565)
Middle	-0.124	-0.131	-0.160
	(0.223)	(0.329)	(0.363)
Rather Left Wing	0.00195	0.0813	0.157
	(0.218)	(0.345)	(0.340)
N	359	179	176

** p < 0.10, ** p < 0.05, *** p < 0.01*

6. Conclusion and discussion

In this study, we used a laboratory experiment to examine preferences for competition when individuals have ability to choose their preferred task. Additionally, we have several treatments that alter the relative information individuals receive about their performance. We found that consistent with our expectations, men generally preferred the stereotypically male task (math) and women preferred the stereotypically female task (word); however, there were many men (women) choosing the word task (math task). Consistent with these choices, men performed somewhat better in the math task and women in the word task. For women, the significant difference in performance over men in the word task only came after they received their preferred (word) task. We conclude that women may have a stronger reaction to this information and increase their motivation relative to men. However, additional studies will need to evaluate this claim more carefully. Relative performance feedback in treatment 2 closes the gender gap by 47 percent when compared to treatment 1 (a gender gap of 31.67 percent vs. 16.67 percent). When the feedback includes relative performance and additional gender information, we find that there is no longer a gender gap. With this respect we argue that even though workers can endogenously choose work task and industries, the gender gap is still large and persistence in our baseline environment. Feedback on one's relative performance is a sufficient mechanism to reduce the gender gap but only when this information includes the performance and of the other participants.

With regards to the competitive entry choice, the choice of task does nothing to change the gender gap compared to previous studies. We find greater than a 2:1 ratio of men to women entering the tournament in treatment 1, which is consistent with the bulk of the literature; we believe this is an important result because of the difference in the experimental environment. To our knowledge, this is the first study to evaluate preferences for competition when individuals can make a choice on their preferred task, and then make a tournament entry decision after they have been sorted into one of two tasks.

Feedback was also useful to induce individuals to state more accurate beliefs about their performance ability. Without any relative performance, feedback men were much more overconfident than women. Interestingly, when individuals received relative performance feedback but did not know the gender of the other competitors' men were still significantly more overconfident than women. Only with additional gender information was there no significant difference in the proportion of overconfident between men and women. Providing feedback on relative per-

formance and gender also increased the percentage of men and women who undertook efficient decisions, that is entering competition when they should or not entering competition when they should not. We find that some sociodemographic variables are significant predictors of the likelihood of men and women entering competition. When both parents work full-time, women are much more likely to enter competition. We find some effect of parental education on the likelihood of entering competition, but only when one of the parents has a Ph.D. Political attitude has no effect on the likelihood of entering competition.

We believe that these results contribute to the literature in several aspects. First, having individuals choose in which task they would like to compete more actually reflects the labor market. The fact that we still find such a large gender gap in competition in our baseline treatment shows if workers can endogenously choose their task (or type of job) in the labor market, this alone cannot alleviate the gender gap. Second, as being able to choose your preferred task more accurately reflects the labor market, we evaluate two information feedback conditions after individuals have made their task preference selection. This is different form earlier studies on information feedback and competitive preferences as we implement this feedback in an environment where individuals can select their preferred task. Our results indicate that firms that want to have more women apply for promotions internally, providing performance feedback only increases the likelihood of women entering competition when this information is couples with additional gender information on other competitors. This is true even when individuals can indicate their preferred task. Finally, when relative performance and gender is provided, workers undertake more efficient decisions – which is beneficial to both workers and firms. Such policies may be useful for firms or policy makers especially when affirmative action programs are difficult to implement or already in place.

References

Balafoutas, L., Kerschbamer, R., & Sutter, M. (2012). Distributional preferences and competitive behavior. *Journal of Economic Behavior and Organization*, *83*(1), 125–135.

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

Black, S. E., & Strahan, P. E. (2001). The division of spoils: rent-sharing and discrimination in a regulated industry. *American Economic Review*, *91*(4), 814–831.

Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, *55*(3), 789–865.

Bock, O., Nicklisch, A., & Baetge, I. (2014). Hroot: Hamburg registration and organisation online tool. *European Economic Review*, *71*, 117–120.

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

Cason, T. N., Masters, W. A., & Sheremeta, R. M. (2010). Entry into Winner-Take-All and Proportional-Prize Contests : An Experimental Study. *ESI Working Paper 10-09.*

Crosetto, P., & Filippin, A. (2013). The “bomb” risk elicitation task. *Journal of Risk and Uncertainty*, *47*(1), 31–65.

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

Dohmen, T., & Falk, A. (2011). Performance Pay and Multidimensional Sorting: Productivity,Preferences,and Gender. *American Economic Review*, *101*(April), 556–590.

Dreber, A., von Essen, E., & Ranehill, E. (2014). Gender and competition in adolescence: task matters. *Experimental Economics*, *17*(1), 154–172.

Eckel, C. C., & Grossman, P. J. (2008). Differences in the Economics Decisions of Men and Women: Experimental Evidence. *Handbook of Experimental Economics Results, Volume 1*, *1*, 509–519.

Ertac, S., & Szentes, B. (2011). The effect of information on gender differences in competitiveness: experimental evidence.

Ewers, M. (2012). Information and Competition Entry. *IZA Discussion Paper*, *6411*(6411).

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

Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in Competitive Environments: Gender Differences. *The Quarterly Journal of Economics*, *118*(3), 1049–1074.

Gneezy, U., & Rustichini, A. (2004). Gender and Competiton at a Young Age. *American Economic Review*, *94*(2), 377–381.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, *104*(4), 1–30.

Große, N. D., & Riener, G. (2010). *Explaining gender differences in competitiveness: Gender-task stereotype* (Jena economic research papers No. 17). Jena.

Günther, C., Ekinci, N. A., Schwieren, C., & Strobel, M. (2010). Women can't jump?—An experiment on competitive attitudes and stereotype threat. *Journal of Economic Behavior and Organization*, *75*(3), 395–401.

Jeworrek, S. (2016). Competition Entry and Relative Performance Feedback: The Importance of Information Disaggregated by Gender.

Kamas, L., & Preston, A. (2012). The importance of being confident; gender, career choice, and willingness to compete. *Journal of Economic Behavior and Organization*, *83*(1), 82–97.

Kerschbamer, R. (2015). The geometry of distributional preferences and a non-parametric identification approach: The Equality Equivalence Test. *European Economic Review*, *76*, 85–103.

Kleinjans, K. J. (2009). Do gender differences in preferences for competition matter for occupational expectations? *Journal of Economic Psychology*, *30*(5), 701–710.

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

Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics*, *122*(3), 1067–1101.

Niederle, M., & Vesterlund, L. (2007). Do Women Shy Away From Competition? Do Men Compete Too Much? *The Quarterly Journal of Economics*, *122*(3), 1067–1101.

Niederle, M., & Vesterlund, L. (2010). Explaining the Gender Gap in Math Test Scores: The Role of Competition. *Journal of Economic Perspectives*, *24*(2), 129–144.

Shurchkov, O. (2012). Under Pressure : Gender Differences in Output Quality and Quantity under Competition and Time Constraints. *Journal of the European Economic Association*, *10.5*(October), 1189–1213.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*(6), 613–629.

Veldhuizen, R. Van. (2016). Gender Differences in Tournament Choices : Risk Preferences , Overconfidence or Competitiveness? *Social Science Research Center Berlin (WZB)*, (November).

Wozniak, D. (2012). Gender differences in a market with relative performance feedback: Professional tennis players. *Journal of Economic Behavior and Organization*, *83*(1), 158–171.

Wozniak, D., Harbaugh, W. T., & Mayr, U. (2014). The menstrual cycle and performance feedback alter gender differences in competitive choices. *Journal of Labor Economics*, *32*(1), 161–198.

This paper investigates the role of feedback in inducing men and women to enter competition. Literature suggests that the underrepresentation of women in management roles and the ensuing gender pay gap might be attributed to differences in preferences for competition, confidence, or risk aversion. A laboratory experiment by the authors found that providing feedback on rank and giving participants the option to compete in their preferred task encourages more women to enter competition. Furthermore, feedback affects reported beliefs and improves efficiency. This study thus provides evidence that manipulating the type of feedback given might be an effective way of closing the gender gap.

ISSN 2509-2359