Do Women Shy Away from Competition?

Muriel Niederle and Lise Vesterlund

April 12, 2005

Abstract

Competitive high ranking positions are still largely occupied by men, and in academia women remain scarce in engineering and sciences. Suggested explanations for this fact focus mostly on discrimination and differences in abilities or preferences (in terms of work hours or field of study). We explore an additional factor, namely that women and men may differ in their selection into competitive environments. In a laboratory experiment, we examine an environment in which women and men perform equally well under both a non-competitive piece rate and a competitive tournament scheme. Participants are then asked to choose the incentive scheme for their next performance. We find that twice as many men as women choose the tournament over the piece rate. This gender gap in tournament entry can neither be explained by performance before nor after the entry decision has been made. While men are more optimistic about their relative performance than women, this difference can only explain a small share of the gender gap in tournament entry. In a final treatment, we show that gender differences exist even when participants simply decide how to be paid for a past performance. We use this decision as a control for non-tournament specific gender differences (such as risk aversion, feedback aversion, general overconfidence), and find a large residual gender effect when participants select tournament compensation for a future performance.

Muriel Niederle: Department of Economics, Stanford University and NBER, http://www.stanford.edu/~niederle, and Lise Vesterlund: Department of Economics, University of Pittsburgh, http://www.pitt.edu/~vester. We thank Scott Kinross, who conducted all the experiments reported in this paper, for his excellent research assistance. We also thank Liran Einav, Jean Francois Richard and Carmit Segal for comments, and we are grateful to the NSF for generous support.

1. Introduction

Gender differences in the representation in high profile jobs remain substantial. For example, Bertrand and Hallock (2001) find that only 2.5 percent of the highest five paid executives in a large data set of U.S. firms are women (for a recent review on the gender differences in wages in the 1990's see Blau and Kahn (2004)). In academia, women are not well represented in science and engineering, and the difference increases with increases in academic rank. In economics there also seems to be a leakage of women at the career ladder (National Science Foundation (2004) and Ginther (2004)). These facts and its causes continue to be the subject of sometimes very heated debates.

Possible (standard economic) explanations for causes of these facts include self selection and discrimination. Women may not select into competitive high profile jobs, or science, economics and engineering, because of the long work hours. Due to intrinsic preferences or intrahousehold bargaining women may prefer to spend more time raising children. Women may also simply not enjoy working in sciences and engineering, or being managers and having many responsibilities. Alternatively, it could be that women have lower abilities (or be less represented among top scorers in ability), which may account for the lack of women in many professions and higher ranks (for a discussion on self-selection see Polachek (1981) and Rosen (1986)). However there seems to be no consensus on the extent to which these (potential) differences in abilities do or would translate into a gender gap. An alternative explanation is discrimination, which leads to differential treatments of women and men with equal abilities and preferences (Black and Strahan (2001); Goldin and Rouse (2000) and see also Altonji and Blank (1999) and references therein.)

In this paper, we want to explore an additional explanation, namely that women may shy away from competitive environments. This not only reduces the number of women who enter tournaments, but also those who win tournaments. Hence it reduces the chances of women succeeding in a competition for promotions, more lucrative jobs, etc. This hypothesis resembles explanations that rely on pure preferences of women. However, a difference is that preferences for tournaments may be malleable, in that they may depend on perceived ability, or general confidence and feelings of competence. The psychology literature suggests that a lack of perceived competence and a fear to receiving negative signals about ones ability (and performance) may be an issue especially in areas in which there exists a stereotype that questions one's ability in the domain.[1] This would be the case for women who enter male-dominated domains, such as managerial positions, and fields such as engineering and sciences. Stereotype threat is one of the bases for the recent wave of support of single-sex schooling or single-sex classes within coed schools (see e.g. Jan (2005)).[2] The concern is that girls may shy away from higher level science classes for fear of competing against boys. The argument for single-sex education is that, maybe, once girls know how able they are, they may be willing to enter competitions against men and be successful at doing so. Surprisingly, however, there has been no experimental test of this assertion, and while there is a vast literature on single-sex schooling and its possible effects, the major obstacle of self selection (of children, or teachers) cannot easily be circumvented.

To test whether women shy away from competition, we use controlled experiments. They allow us to measure performance precisely, and to exclude any discrimination or expectation of discrimination on the part of participants. We have groups of 2 women and 2 men perform a real task, namely adding up five two-digit numbers for five minutes. In order to assess possible gender differences in the choice of the incentive scheme (i.e., a competitive or a non-competitive payment scheme) we opt for a task in which we expect no gender differences in performance under either of these incentive schemes.[3] Participants first perform in a piece rate scheme and then in a tournament scheme. Indeed we observe no gender differences in performance under either incentive scheme. Participants do not receive any feedback concerning their performance, but experience both incentive schemes.

We then have participants choose the incentive scheme for their next performance, either a piece rate or a tournament. The competitive incentive scheme is designed in a way that reduces the participants' problem to an individual decision making problem, such that the participant does not have to form beliefs over other players' choices of incentive scheme. We find that twice as many men as women enter the tournament. This gender difference cannot be explained by past performances. In general, the entry decision is not very correlated with the participants' past tournament performance. Furthermore, the participant's performance after the decision is also a poor predictor of the entry decision, and it fails to account for the gender gap in tournament entry.

[1] For the literature on stereotype threat theory and its implications see Steele (1997).
[2] For possible benefits of single sex schooling see also Harwarth, Maline and DeBra (1997) and Solnick (1995).
[3] Gneezy, Niederle and Rustichini (2003) showed that this may not be straightforward. We will tie our results to the past literature in the last section of the paper.

Since participants are not aware of their relative performance, their entry decisions may be driven by their beliefs about their relative performance ranking. Our study shows that while both genders are overconfident about their relative ranking, men are significantly more optimistic than women. However, this gender difference in believed ranking can only account for a small share of the gender difference in tournament entry.

There are several reasons why women may shy away from competitions. While some are specific to tournaments and performing under this compensation, others are not. The non-tournament specific factors may include that women are more risk averse than men, and that they are more averse to receiving feedback on their relative performance. Women may also consistently hold less optimistic beliefs than men, and may not be very confident in the point predictions of their beliefs. We present the participants with one last task to assess whether these explanations by themselves generate a gender gap in entry decisions, or if tournament-specific explanations and future performance are needed. We also use this task to assess how much of the gender gap in tournament entry can be accounted for by explanations that are not specific to tournaments, and how much is due to the fact that participants enter a tournament and have to perform once more. Tournament specific explanations include, for example, that women may have psychic costs of participating in a tournament (or of course, men could receive psychic benefits), or that women may be more averse to receiving feedback on a tournament performance. It could also be that women are especially under confident and uncertain in their tournament performance, or that women feel that a good past performance is a poor predictor for a good future performance (see e.g. Beyer (1990) and Felder et al (1994)).

In this last task, participants are compensated once again for their past non-competitive piece rate performance, but now they select the compensation scheme they want applied to their performance. Participants can decide to be paid according to the piece rate, or they can submit the past piece rate performance to a tournament. They win the tournament, if their past piece rate performance is the highest in their group. Once again we find a significant gender gap in the decision to submit the piece rate to a tournament, and this difference cannot be explained by the piece rate performance. We also find that while both men and women are overconfident about their relative performance, men are more optimistic than women. The difference relative to the initial tournament entry decision is that when submitting the piece rate gender differences in beliefs can account for the gender gap in selecting the tournament:

Controlling for the formed beliefs, there is no significant gender difference in the rate of submitting the past piece rate result to a tournament.

Behavior in this last task suggests that in this experiment gender differences in submitting the past piece rate performance to a tournament are mostly driven by gender differences in beliefs on relative rank.

We then use behavior in this final task as a control for gender differences that may result from non-tournament specific explanations, and try to estimate the existence and size of the residual gender difference in the decision to enter a tournament and subsequently perform. We find that gender differences are still significant and large. That is, a sizable portion of the gender gap in tournament entry seems to be driven by gender differences that are specific to the fact that participants perform once more in a tournament.

2. Experimental Design

We want to establish an environment in which we can measure the propensity by which women and men choose a competitive tournament incentive scheme over a noncompetitive piece rate scheme. We also want to examine to what extent these choices are driven by performance. For this purpose we conduct an experiment in which participants have to solve a real task. We measure the performance of participants under both a piece rate and a tournament incentive scheme, and subsequently observe their choice between performing under these two incentive schemes.

To attribute potential gender differences to preferences over incentive schemes, as opposed to ability differences, we aim for a task in which we do not expect gender differences in performance, under both a piece rate and a tournament scheme.[4]

The task of our experiment is to add up five 2-digit numbers. Participants could not use a calculator, but were given scratch paper to write down numbers if they wished to do so. The numbers were randomly drawn and presented in the following way:

21	35	48	29	83	

[4] This is not necessarily straightforward, see Gneezy, Niederle and Rustichini (2003). At the end of the paper we try to provide a more unified framework on the possible reasons for gender differences and its effects on labor market outcomes.

Once the participant submitted an answer on the computer, a new problem appeared jointly with information on whether the former answer was correct or wrong.[5] Furthermore a record of the correct and wrong answers was kept on the screen. Participants had 5 minutes in which they received as many problems as they wanted to answer.

The task is one in which skill and effort are relevant. We did not expect any gender differences, as adding up numbers is an easy math test, in which there are no gender differences.

The experiment was conducted at the University of Pittsburgh, using the PEEL subject pool and standard recruiting procedures. Two or three groups of 4 participants, two women and two men, participated in each session. Although gender was not discussed at any time, participants could see each other in the lab, and determine the gender of the participants in their group. A total of 20 groups participated in the experiment (80 participants). Each participant received a $5 show-up fee, and an additional $7 for completing the experiment. Participants were told that they would be asked to complete four tasks, and that one of these tasks would be randomly chosen for payment at the end of the experiment. By paying only one task, we avoid problems that the decisions in any given task may be influenced by insurance or hedging motives given the outcomes of other tasks. Subjects were informed of the nature of the tasks only immediately before performing the task. While subjects know how well they do on a task, i.e. how many problems they solve correctly, they are not informed of their relative performance until the end of the experiment. Participants first perform in both a noncompetitive piece rate scheme and a competitive tournament scheme. All participants therefore experience both incentive schemes, and we have a performance measure of each participant under each incentive scheme.

Task 1- Piece Rate:

Participants are given five minutes in which to calculate sums of series of five randomly chosen two-digit numbers. If task 1 is selected for payment, subjects receive 50 cents per correctly solved problem. Participants will not be informed about the number of problems solved by other participants.

[5] The program was written using the software zTree (Fischbacher 1999).

Task 2 – Tournament:

Participants are given five minutes in which to calculate sums of series of five randomly chosen two-digit numbers. If task 2 is selected for payment, the participant who solves the largest number of problems in the group receives $2 per correct problem, while the other participants receive no payment (in case of ties the winner is chosen randomly). Participants receive no feedback whether they won the tournament, or how many problems were solved by others.

In the third task, we give each participant the choice whether to perform in a piece rate or in a tournament. We want the choice of each participant to depend on their preferences for different incentive schemes and their beliefs regarding their relative performance in a tournament scheme. We want the choice of a participant to be independent of the participants' beliefs about others' choices, therefore we eliminated any externality that the choice of one participant imposes on another one. This allows us to exclude that women may shy away from competition because they worry of imposing a negative externality on men (e.g., by winning the tournament.)[6] It also means that participants do not have to estimate which other participant may enter the tournament. We thereby avoid an additional source of errors through wrong beliefs about other participants choices. Participants therefore face an individual decision making problem which depends only on their beliefs about their potential to win their tournament and their preference for performing in a tournament.[7]

Task 3 – Choice:

Participants are given five minutes to calculate sums of series of five randomly chosen two-digit numbers. They can choose whether they want to be paid according to a piece rate, i.e., 50 cents for each correctly solved problem, or a tournament. When the participant chooses tournament, her new performance will be evaluated against the task-2 tournament performance of the other participants in her group. If the participant has the highest performance she receives $2 for each correct answer, otherwise she receives no payment (in case of ties the

---

[6] For a discussion on possible gender differences in altruism see Andreoni and Vesterlund (2001), and see Ledyard 1995 for gender differences in dilemma and public good games.

[7] By reducing the problem to an individual decision problem, we do not need to assess the participants beliefs about other participants' choices, which would be crucial would we try to understand behavior in a game in which the decision were to depend on other players choices. For example, it could be that all men believe no woman enters a tournament, meaning that they face only one other competitor, greatly changing the odds to win the competition.

winner is chosen randomly). Participants will not be informed of how they did in the tournament until all tasks of the experiment have been completed.

We chose to compare the performance of participants who choose tournament in task 3 with the task-2 tournament performance of the other participants in her group as opposed to the other participants' task-3 performance. One can think of this as competing against other participants who already performed.[8] This has two advantages; first, the performance of a player who enters the tournament is evaluated against the performance of participants that performed under the same tournament incentive scheme. It also implies that a participant who enters the tournament has to think only about her beliefs regarding the other players' tournament performance, and not whether other players are more or less prone to enter the tournament and therefore may have a higher or lower performance. A participants' choice does not affect the payment of any other participant in any way.[9]

Once we established gender differences in choosing the tournament, we try to understand why women shy away from competition. It could be that women simply do not like to perform in a competitive environment, or that women simply shy away from environments with uncertain payments. It could also be that women are less confident in their performance, or less confident in being able to judge their performance. These are a few of many factors that can drive a gender difference. In the last treatment we want to assess whether there are still gender differences, even when we eliminate all aspects that strictly relate to a future tournament performance. We also want to estimate whether women shy away from tournaments because of the competitive nature, or for reasons that are not necessarily only associated with tournaments (such as higher levels of risk aversion). Therefore, we try to subject women and men to the same kind of choice as in task 3, however without having participants actually perform in a tournament and without using any tournament performance. This allows us to eliminate all aspects of the choice that are specifically related to performances in a competitive environment, while keeping all other aspects. Specifically, it allows us to eliminate whether women simply do not want to perform in competitive

[8] Many sports competitions are organized that way, for example downhill skiing, and generally all competitions in overall rankings are competitions where participants do not necessarily all perform simultaneously.

[9] Our design allows for the possibility that there is no winner among participants who choose the tournament (if, e.g., the highest performing participant opts for the piece rate). Conversely, all participants can win the tournament, if everyone increases their performance beyond the highest task-2 performance in that group. That is, a participant imposes no negative externality on other participants by choosing the tournament option.

environments, or that women may have beliefs about performing less well in competitive environments. However, the decision will include other aspects that are not directly related exclusively to performing in a tournament, such as e.g. the fact that the choice is between a certain and an uncertain payment. That is, in the next treatment participants face a choice similar to the one in task 3, without using the tournament performance, and without having participants to perform in a tournament.

Task 4 – Submit Piece Rate:

In the final task, participants do not have to perform rather they will be evaluated according to the number of problems they solved in the task-1 piece rate. Participants can choose how to be paid in the 4th task. They can decide to be paid by a 50 cent piece rate, or according to a tournament. They win the tournament, if their task-1 piece-rate performance is the highest among all the participants in their group, in which case they receive $2 for each correct answer, otherwise they receive no payment (in case of ties the winner is chosen randomly). Before making their choice, participants are reminded of their task 1 piece rate performance.

We chose to pay participants according to the task 1 piece rate, as opposed the task 2 tournament, to eliminate all aspects of performing in a tournament in the 4th task.[10]

This final task will allow us to see whether gender differences appear even in a choice that does not involve a future tournament performance. Furthermore, we will use the decision in this last treatment as a control for the effect of uncertain payment, possible gender differences in assessing and acting upon one's beliefs in relative performance, and so on, that is all effects that are present in the task 3 choice, but are not specific to tournaments only. With this decision as a control we estimate whether there is an additional gender difference when it comes to choosing to enter a competitive environment.

The participants' decision to enter the tournament or submit the piece rate result to a tournament may be driven by absolute performance; however payments depend on relative performance. Participants are not aware of the relative ranking of their performance, since they receive no feedback during the experiment. Participants have to form beliefs about their

[10] As in the choice in task 3, a participants decision in task 4 does not affect any other participants' earnings and is hence an individual decision making task.

relative performance. Their decisions depend on their beliefs, which may or not be accurate given their performance. To be able to evaluate the effects of beliefs, and how participants form beliefs as a function of their performance and the payment scheme, we elicit the beliefs of any performance that is relevant for the decisions of participants.

Final Questions:

At the end of the experiment, we ask participants to guess their rank in the task-1 piece rate and the task-2 tournament. They had to pick one of the numbers 1,2,3,4, and were paid $1 for each correct answer.[11]

At the end of the experiment, a random number from 1 to 4 was drawn, which determined the task which was relevant for earnings. The experiment lasted about 45 minutes, and participants earned on average $ 20.

3. No Gender Differences in Performance

The objective of this study is to examine whether, conditional on ability, men and women differ in their preference for performing under a piece-rate versus a tournament scheme. We want male and female participants to be as similar as possible in their performance under both a piece rate and a tournament compensation, in order to eliminate ability differences as a possible explanation for any potential gender differences when choosing the environment in which to perform. We first confirm that we found such a task in which men and women are equally productive under the two compensation schemes.

Under piece-rate compensation there is no significant gender difference in performance. The average number of problems solved is 10.15 for women and 10.68 for men, there is no significant difference.[12] Figure 1 shows the cumulative distribution of the number of solved problems for women and men.

[11] In case of ties in the ranks, we counted every answer that could be correct as correct. For example, if the performance in the group was 10, 10, 11, 11, then an answer of last and third was correct for 10, and an answer of best and second was correct for 11.

[12] The standard error is 0.44 for women and 0.55 for men. A two-sided Mann Whitney yields p=0.62 and a t-test yields p=0.459. The average number of incorrect answers is 2.8 for both men and women. Mean earnings are $5.3 for men and 5.08 for women.

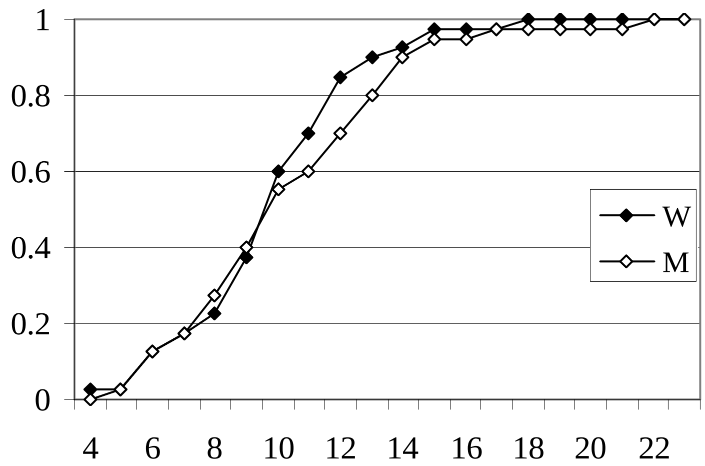

Figure 1: CDF of Piece-Rate Performance (Task 1)

The average number of problems solved in the tournament is 11.8 for women and 12.1 for men. Once again there is no significant gender difference in mean performance.[13] Figure 2 shows that there is no gender difference in the distributions of number of problems correctly solved.

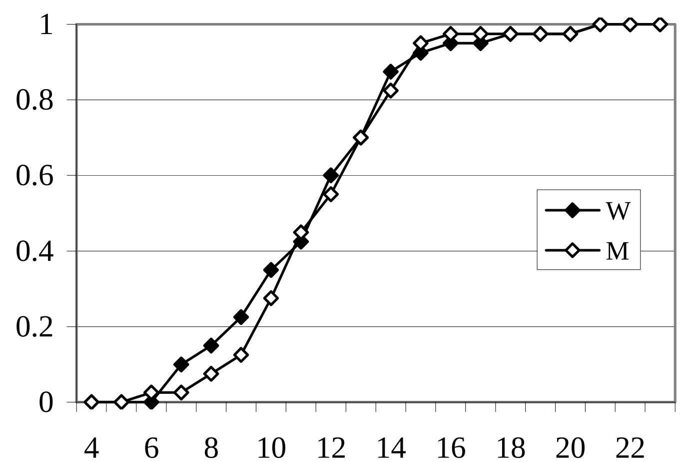

Figure 2: CDF of Tournament Performance (Task 2)

[13] The standard error is 0.48 for women and 0.43 for men. A Mann Whitney yields p=0.594, and a t-test p =0.643. The average number of incorrect answers is 3.0 for men and 2.3 for women. A t-test reveals that this difference is significant with a p-value of 0.033.

The performance of both men and women is significantly higher under tournament than piece-rate compensation.[14] This can be caused by learning or by the different performance incentives under the tournament.[15] The increase in performance varies substantially across participants, while this may simply be noise, it could also be due to the fact that some participants are more competitive than others. Note however that the increase in performance from the piece rate to the tournament does not differ by gender.[16] Furthermore, the piece rate and tournament performances are very correlated, with spearman rank correlations of 0.69 for women and 0.61 for men.

The similar performance of men and women is reflected in their similar probabilities of winning the task 2 tournament. Of the 20 groups in our task-2 tournament, 11 were won by women and 9 by men.[17] To assess probabilities of winning the tournament, we randomly create four-person groups from the observed performance distributions. Conditioning only on gender the probability of winning the tournament is 26% for a man and 24% for a woman, and as shown by Table 1 the probability of winning conditional on performance is also the same across gender.[18]

Table 1: Probability of Winning Task-2 Tournament Conditional on Performance

	8	9	10	11	12	13	14	15	16	18	21
Women	0.1	0.4	1.8	5.5	13.4	26.6	47.8	71.9	84.6	90.3	96.3
Men	0.1	0.4	1.8	5.6	13.4	26.6	47.7	71.9	84.6	--	96.3

We therefore succeeded in selecting a task for which there is no gender difference in either a piece rate and a tournament performance. Hence after having completed the first two tasks women and men have similar experiences and on the basis of performance we would not

[14] The increase in number of problems solved under the tournament is 1.68 for women, and 1.45 for men (std.err. of 0.36 and 0.39 respectively). A one-sided t-test yields p<0.01 for each gender separately.

[15] Malmendier, Della-Vigna, and Vesterlund (2005) have participants perform in series of 3 minutes tournaments, and find a significant increase in performance from round 1 to round 2, but no significant increase in performance in subsequent rounds. This suggests that learning how to perform in this task may have some effect.

[16] A Mann Whitney test yields a p-value of 0.553, and a t-test p=0.673. In our sample, 37% of men and 25% of women experience no or negative change.

[17] The 11 female winners had performances of 12, 13, 14, 14, 14, 14, 15, 15, 16, 18, 21, and hence an average of 15.09. The performance of the 9 men is 12, 14, 14, 15, 15, 15, 15, 16, 21, with an average of 15.22. These differences result in women earning an average of $8.3 and men $6.9, this difference is not significant p=0.63.

[18] For any given performance level, say 15 for a woman, we draw 10,000 groups consisting of 2 men and one other woman, where we use the sample of 40 men and women with replacement. We then calculate the frequency of wins. The exercise is repeated 100 times and we report the average of these win frequencies.

expect any gender differences when participants choose between the two different compensations for the third task.

4. Do Women Shy Away from Competition?

After having performed in the piece rate and in the tournament participants are asked whether they want the compensation for their task-3 performance to be a piece rate or a tournament. The payment per problem solved is 50 cents under the piece rate, and $2 when winning the tournament. A participant who chooses the tournament wins the tournament if the number of problems she solves in task-3 exceeds the number of problems solved in task-2 by the other three members of her group. Therefore, choosing the tournament depends on beliefs regarding the other players' past tournament performance, but not on the player's belief about the task-3 compensation choice of the other participants.

Ignoring the costs of performing the task, a risk-neutral participant is indifferent between the two incentive schemes, when her chance of winning the tournament is 25%. If the participant's task-3 performance is exactly like her task-2 performance, then 30% of the women and 30% of the men (all those who have a performance of 14 and higher) would have higher expected earnings from a tournament than a piece-rate scheme. When we include participants who solve 13 problems – and are virtually indifferent between the two incentive schemes – the percentages are 40% for women and 45% for men.

Despite the identical performances in the piece rate and tournament, women and men differ in their choice of compensation. While the majority of women prefer the piece rate, the majority of men prefer the tournament: 35% of women and 73% of men select the tournament. This difference is statistically significant.[19]

5. Does Performance Predict the Compensation Choice?

We start our investigation of the gender difference in compensation choice by examining whether it can be explained by past and future performance, that is, performance before and after the task-3 compensation choice.

5.1. Does Past Performance Predict Tournament Entry?

To investigate if past performance is a good predictor of compensation choice we first compare the mean performance characteristics of participants who chose piece-rate to those

[19] A chi square test yields a p-value of 0.001, and a fisher's exact test a p-value of 0.002.

who selected a tournament compensation. Table 2 reports, by gender and the chosen compensation, the average number of problems solved under piece rate and tournament, as well as the average increase in performance between the two.

Table 2: Performance Characteristics by Compensation Choice

		Average Performance		
	Compensation Choice	Piece Rate	Tournament	Tournament – Piece Rate
Women	Piece Rate	10.35	11.77	1.42
		(0.61)	(0.67)	(0.47)
	Tournament	9.79	11.93	2.14
		(0.58)	(0.63)	(0.54)
Men	Piece Rate	9.91	11.09	1.18
		(0.84)	(0.85)	(0.60)
	Tournament	10.97	12.52	1.55
		(0.67)	(0.48)	(0.49)

Notes to table: standard errors in parenthesis

For women there is no difference in performance between those who chose the tournament and those that do not. Mann-Whitney tests on the performance differences between women who select the piece rate and those that enter the tournament yield p-values higher than 0.23.[20] A probit regression of the decision to enter the tournament on past performance reveals that the coefficients on tournament performance and the increase in performance from the piece rate to the tournament are not significantly different from zero.[21]

For men, only the tournament performance seems to be marginally higher for those who select the tournament for task 3.[22] A probit regression of the decision to enter the tournament on past tournament performance and the increase in performance from the piece rate shows a marginal effect of the task-2 tournament performance of 0.04 (s.e. 0.02, p=

[20] Comparing the performance of the women who select the tournament and those that do not yields the following test statistics. Mann-Whitney tests yield: p=0.44 for task-1 piece rate, p=0.51 for task-2 tournament, and p=0.23 for the difference in performance between task 2 and task 1. t-tests yield: p=0.55 for task-1 piece rate, p=0.88 for task-2 tournament, and p=0.35 for the difference in performance between task 2 and task 1.

[21] A probit analysis of women's compensation choice on task-2 performance and the performance difference between task 2 and 1 yields coefficients on tournament performance of -0.01 (s.e. 0.03, p=0.71), on difference in performance of 0.04 (s.e.0.04, p=0.31), when we evaluate the marginal effects at a tournament performance of 13 and a piece rate performance of 12.

[22] Comparing the performance of men who select the tournament and those who do not yields the following test statistics: Mann-Whitney tests yield p=0.46 for task-1 piece rate, p=0.23 for task-2 tournament, and p=0.51 for the difference in performance between task 2 and task 1. t-tests yield: p=0.40 for task-1 piece rate, p=0.14 for task-2 tournament, and p=0.68 for the difference in performance between task 2 and task 1.

0.15).[23] That is, tournament performance has only a marginal, very weak effect on the entry decision.

Figure 3 shows the proportion of women and men who enter the tournament conditional on their performance quartile: For every performance level men are more likely to enter the tournament. Even women in the highest performance quartile have a lower propensity to enter the tournament than men in the lowest performing quartile.

Furthermore, among high performing participants, i.e. those who solve 13 and more (and hence whose past performance would yield higher expected earnings in a tournament than a piece rate) significantly more men than women enter the tournament. Similarly there is a significant gender difference in the entry decision among low performing participants.[24]

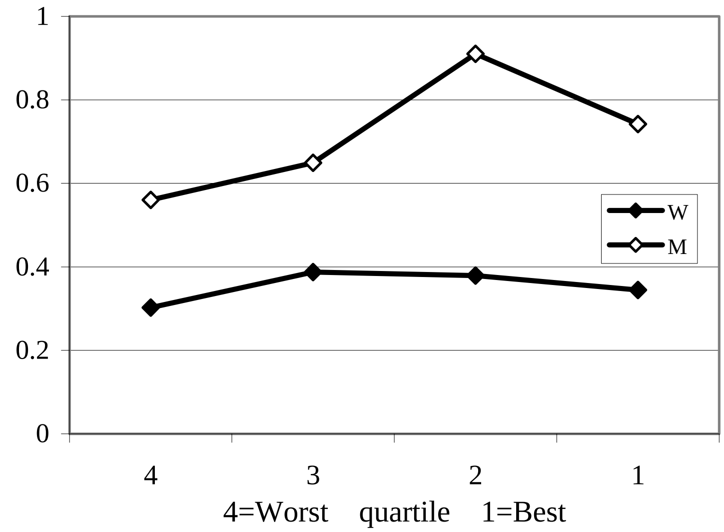

Figure 3: Proportion of participants entering the tournament conditional on the Task-2 tournament performance quartile[25]

The observation that, conditional on performance, men enter the tournament more than women is confirmed by a probit regression, see Table 3.

[23] A probit of men's compensation choice on task-2 performance and the performance difference between task 2 and 1 yields coefficients on tournament performance of 0.04 (s.e. 0.02, p=0.15), on the difference in performance of 0.01(s.e. 0.03 p= 0.83), when we evaluate the marginal effects at a tournament performance of 13 and a piece rate performance of 12.

[24] Among participants with a task 2 tournament performance of 13 and higher, 15/18 (83%) of men enter the tournament compared to 5/16 (31%) of women, this difference is significant, a Chi square test delivers p=0.002. The difference is also significant at a 4% level when we include only participants who solve 14 or more. For participants who solve 12 or less in task 2, 14/22 (64%) of men compared to 9/24 (38%) of women enter the tournament, a Chi square test delivers p=0.08.

[25] For each performance quartile we report the propensity by which men and women in this quartile enter the tournament. Since there are no gender differences in performance in this task there are approximately equal numbers of men and women in each quartile.

Table 3: Probit of Tournament Choice on Individual Characteristics

	Coefficient	Standard Error	p-value
Female	-0.381	0.104	0.01
Tournament	0.018	0.022	0.41
Tournament – Piece Rate	0.018	0.027	0.50

Notes to table: Dependent variable: task-3 compensation choice (1-tournament and 0-piece rate). Independent variables: a female dummy, task-2 tournament performance, change in performance between task-2 and task-1. The table presents marginal effects on the probability of selecting the tournament evaluated at a male solving 13 problems in the tournament and 12 in the piece rate.

The decision to enter the tournament is not strongly related to the participants' performance under the two compensation schemes, though men with a higher tournament performance are somewhat more likely to select this compensation. However, men and women who select the tournament performed the same under both piece rate and tournament, and similarly there is no gender difference in ability among those who select the piece rate.[26]

Thus controlling for ability women are much less likely to select to perform in a competitive compensation scheme than any of their male equivalents.

One possible explanation of the observed gender difference in preferred compensation may be that there is a gender difference in performance following the compensation choice – and that our participants correctly anticipate such a difference.

5.2: Does Future Performance Predict Tournament Entry?

To assess if task-3 performance can account for the decision to enter the tournament, we compare, for each gender, the performance conditional on the participant's task-3 compensation choice.

[26] The p-values when comparing the performance of women to men who did enter the tournament, range from 0.28 to 0.57 when using Mann Whitney or t-tests. Similarly, when comparing women to men among participants who did not enter the tournament (p-values from 0.47 to 0.77). There are also no differences in the number of incorrect answers. The average number of wrong answers under the piece rate is 2.5 for men and 2.8 for women, in the tournament the number is 2.9 for men and 2.7 for women. Neither of these differences are significant (p=0.517 and 0.794 respectively).

Table 4: Choice Performance Conditional on Compensation Choice

	Compensation Choice	Average Performance: Choice	Average Performance: Choice-Tournament
Women	Piece Rate	11.62	-0.15
		(0.62)	(0.44)
	Tournament	11.79	-0.14
		(0.64)	(0.50)
Men	Piece Rate	11.91	0.82
		(0.91)	(0.44)
	Tournament	13.48	0.97
		(0.83)	(0.55)

Notes to table: standard errors in parenthesis

For each gender, there is no significant difference in task-3 performance between those who do and do not enter the tournament.[27] Nor do participants who enter the tournament have a significantly different increase in performance in the task 3 choice treatment compared to the tournament treatment.[28]

That is, not only is it not true that only participants with a high past performance enter the tournament, it is also not true that those that entered the tournament performed better than those that did not.[29]

While, for a given gender, the task-3 performance does not predict which participants enter the tournament, it may account for the gender differences in deciding whether to enter the tournament or not. Figure 4 shows the cumulative distribution of the performance of women and men in task 3, independent of the chosen compensation scheme.

[27] A Mann-Whitney test yields p=0.63 for women and p=0.24 for men (a t-test respectively p=0.861 and p=0.288).

[28] A Mann-Whitney test yields p=0.93 for women and p=0.58 for men (respective p-values from a t-test are 0.99 for women and 0.88 for men).

[29] A probit analysis of the decision to enter as a function of the task 3 poerformance yields marginal effects of a participant that solves 13 in the task 3 treatment of 0.02 (s.e. 0.02, p=0.265) for men and 0.005 (s.e. 0.03, p=0.85) for women. That is, there is no significant effect of the performance on the decision to enter the tournament.

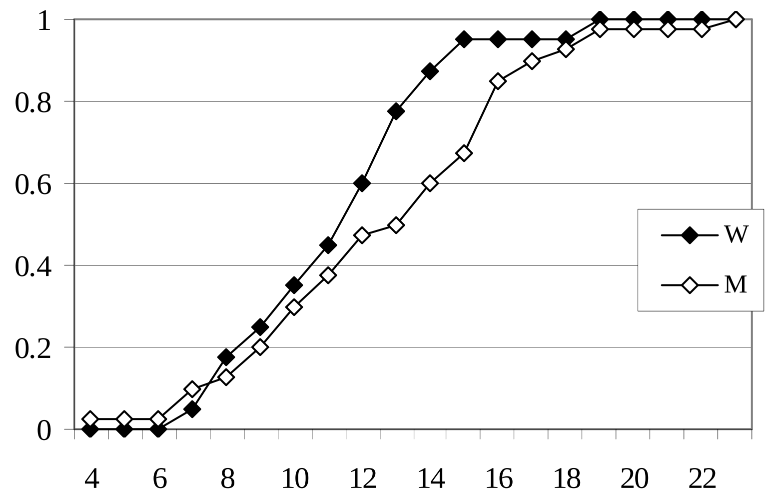

Figure 4: The cumulative distribution function of Choice Performance (Task-3).

While men solve an average of 13.05 problems in task 3, women solve only 11.68 problems. In contrast to our earlier results we observe a significant gender difference in the task-3 performance.[30] Compared to the task-2 tournament, men significantly improve their performance solving on average one additional problem in task 3, whereas women do not change their performance.[31] Furthermore the improvement in performance for men is significantly larger than it is for women.[32]

However, this difference in performance *cannot* account for the gender difference in the choice under which incentive scheme to perform. We have seen in Figure 3 that even men in the lowest performing quartile of the task-2 tournament (who solve between 6 and 10 problems) have a higher propensity to enter the tournament than women in the top performing quartile (who solve between 13 and 21 addition problems).[33]

To assess the correlation between the actual performance in task 3 with the choice of the incentive scheme, i.e. to assess whether the choice was driven by a correct anticipation of performance, we compute the decision to enter the tournament as a function of the participants' task-3 performance. Figure 5 shows the proportion of participants that chose tournament for each performance quartile. The gender difference in the tournament entry decision conditional

[30] The standard error in task 3 performance of men is 0.65 and of women 0.45. A Mann-Whitney test comparing the performance of women and men yields p = 0.082 (t-test p=0.088).
[31] p-values of a t-test for men and women are 0.03 and 0.65, respectively.
[32] A Mann-Whitney test comparing the increase in performance of women and men yields p = 0.03, and a t-test p=0.046.
[33] In fact men in the lowest performance quartile have an average increase in performance of 0.42, compared to -1 for women in the highest performance quartile. These changes are not even sufficient to have the performance intervals overlap.

on task-3 performance is very similar to that observed conditional on the performance in the task-2 tournament. Independent of performance women are much less likely to choose to enter the tournament.

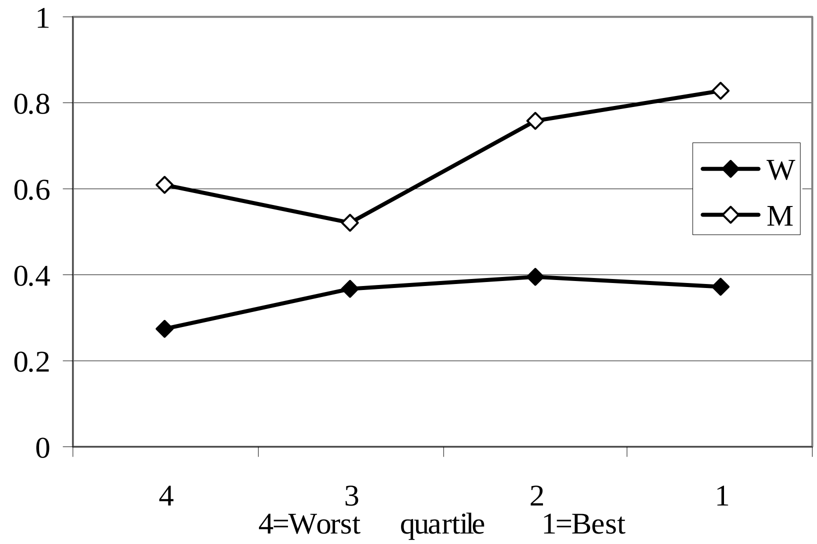

Figure 5: Proportion of Participants Entering the Tournament Conditional on Task-3 Performance Quartile.[34]

We observe that among both groups, namely participants who have a higher than 25% chance to win the tournament, and those that have a lower 25% chance to win the tournament, significantly more men than women enter the tournament.[35]

The probit regression in Table 5 confirms that there is a significant gender difference in the decision to enter the tournament when conditioning on future, task 3 performance. Furthermore, the performance is not correlated with the entry decision. That is, the gender difference in task 3 performance cannot explain the gender difference in tournament entry.

[34] For each performance quartile we report the propensity by which men and women in this quartile enter the tournament.

[35] Among participants who solve 13 and more (i.e. have a higher than 25% chance to win the tournament), 17/21 (81%) of men and 6/16 (38%) of women had decided to enter the tournament, the difference is significant, a chi square test yields p=0.01. Similarly, among participants who have a lower than 25% chance to be the highest performer, 12/19 (63%) and 8/24 (33%) of women had chosen tournament, this difference is significant, a Chi-square test yields p=0.05.

Table 5: Probit of Task-3 Compensation Conditional on Performance

	Coefficient	Standard Error	p-value
Female	-0.359	0.106	0.002
Task-3 performance	0.018	0.017	0.311

Notes to table: Dependent variable: task-3 compensation choice (1-tournament and 0-piece rate). Independent variables: female dummy and performance in task 3. The table presents marginal effects on the probability of selecting the tournament evaluated at a male solving 13 problems in task-3.

That is, even when controlling for the fact that men increase their performance in the third task (while women do not), women behave very differently from men and enter the tournament significantly less. There is no strong correlation between compensation choice and past or future performance for either women or men.

5.3 Is the Entry Decision Determined by Beliefs?

Actual performance can not explain the gender differences in choosing the incentive scheme. However, the decision to choose a competitive compensation depends not on an individual's absolute ability, but rather on how the ability ranks relative to the other group members. While participants are aware of their absolute performance, they do not know their relative ranking. Hence the compensation choice is based on their beliefs about their relative performance.

Before we assess the importance of believed rankings on the choice of a tournament scheme, we study the formation of beliefs. Do men and women of equal ability differ in their assessment of their relative performance, and if so, can these potential confidence differences account for the gender difference in compensation choice. That is conditional on beliefs about relative performance do we observe men and women being equally likely to choose the tournament.

To investigate the impact of the believed ranking in the task-2 tournament performance we ask participants at the end of the experiment, to guess how their performance in task-2 ranked relative to the other members of their group. Participants received $1 if their guess was correct, and in the event of tie they were compensated for any guess that could be deemed correct. For example, in the event of a two-way tie for first rank, a guess of either first or second rank would warrant the $1 payment.[36]

5.3.1. Do Women and Men Form the Same Beliefs?

To measure the effects of performance, specifically the tournament performance and the increase between tournament and piece rate on the guessed rank of participants, we use an ordered probit regression.[37] For both men and women we find that a higher tournament performance is significantly associated with a better guessed rank (i.e., a guess of performing better).[38] That is, independent of gender the guesses on rank are positively correlated with actual performance, though only weakly so.

Before we study the actual distribution of guesses, we determine the distribution of perfectly calibrated individuals. Due to the tie-breaking rule participants who have a uniform prior over their position in the performance distribution have an incentive to guess that they are ranked second or third.[39] If the participant is aware of the performance distribution and their own performance then the payoff-maximizing distribution of guesses differ a bit more across gender. In Table 6 we report the actual distribution of ranks as well as the distribution of guessed ranks of perfectly calibrated and payoff maximizing participants.[40]

---

[36] While the payment for the guessed rank is not very high, it may still offer participants the opportunity to use a guess as a potential source of hedging. Our results to be presented clearly indicate that this was not a motive for the majority of participants.

[37] We eliminate guessed ranks of 4, as we have very few data points there. The results are similar when we code guesses of 3 and 4 as guesses of rank 3.

[38] An ordered probit regressing the belief on the tournament performance and the increase in performances yields for women a tournament coefficient of -0.24 (s.e. 0.09, p-value 0.01) and a coefficient on the increase in performance of 0.02 (s.e.0.12, p-value 0.85). For men, the coefficients are -0.17 (s.e.0.098, p-value 0.086) and -0.12 (s.e.0.019, p-value 0.19). The average performance of male participants with a guessed rank of 1 is 12.5, compared to 12 for a guessed rank of 2, 9.5 for a guessed rank of 3. For women, the average performance for a guess rank of 1 is 13.1, for a guess of 2 is 11.7 and for a guess of 3 is 9. Optimal guesses are 1 for a performance of 14 and higher, 2 for a performance of 12 and 13, and 3 for a performance of 10 and 11.

[39] Based on 10,000 artificially generated groups the likelihood of a woman being ranked first is 0.223, second 0.261, third 0.262, and last 0.255, the corresponding probabilities for a man for first is 0.243, second 0.288, third 0.278, and last 0.199.

[40] These distributions are based on the results of 10.000 randomly generated groups for each performance level of men and women. Neither the actual rankings, nor the optimal guesses differ significantly across gender (the p-values of Chi-square tests are 0.54 and 0.50 respectively).

Table 6: Distribution of Actual and Optimal Guess of Tournament Rank

	Men		Women	
	Actual Rank	Optimal Guess	Actual Rank	Optimal Guess
1: Best	9	12	11	12
2	16	10	11	11
3	9	13	8	8
4: Worst	6	5	10	9

Notes to table: Rank out of four. Actual rank is based on the highest rank that could possibly be correct in the case of ties.

As shown in Table 7, the distribution of guessed ranks differs substantially from those reported in Table 6.

Table 7: Distribution of Guessed Tournament Rank

	Men		Women	
	Guessed Rank	Incorrect Guess	Guessed Rank	Incorrect Guess
1: Best	30	22	17	9
2	5	3	15	10
3	4	2	6	5
4: Worst	1	1	2	1
Total	40	28	40	25

Notes to table: Guessed rank out of four.

Even though the tournament performance of men and women is basically identical, men have higher beliefs about their performance, 75% of the men think they are the best in their group of 4, compared to 43% of the women. The guesses of women and men are significantly different, a Chi-square test delivers p=0.025.[41] Table 7 also reports the distribution of wrong guesses, men are more likely to incorrectly guess that they are ranked first. The distributions of incorrect guesses differ significantly across gender.[42] Finally, relative to the optimal and actual guesses both men and women appear overconfident. While the distribution of guesses for men differs significantly from both the optimal and actual distribution, that of women only differs significantly from the actual distribution.[43]

[41] A Mann Whitney test comparing the beliefs of men and women yields a p-value of 0.008.

[42] A Mann Whitney test on the distribution of wrong guesses of men and women delivers a significant difference, p<0.01. The p-value of a Chi-square test is 0.015.

[43] A Chi-square test of independence between the distribution of guessed rank and optimal guessed rank yields p=0.0008 for men and p=0.102 for women, the test of independence between distribution of guessed rank and actual rank yields p=0.0001 for men and p=0.057 for women.

To determine how beliefs relate to performance and whether women and men form different beliefs conditional on performance, we use an ordered probit to estimate the guessed rank as a function of tournament performance, the increase in performance and a female dummy. The results in Table 8 reveal that while both men and women base their guessed rank on their actual performance, women are significantly less optimistic about the ranking of their performance than men.[44]

Table 8: Ordered Probit of Guessed Rank

	Coefficient	Standard Error	p-value
Female	0.75	0.30	0.01
Tournament	-0.19	0.06	0.003
Tournament – Piece Rate	-0.08	0.07	0.27

Notes to table: Ordered probit of guessed rank as a function of performance and a female dummy, for guesses of ranks 1,2 and 3.

Both women and men base their beliefs weakly on performance. However, conditional on performance, men have significantly better beliefs about themselves than women.

5.3.2. Do Beliefs Predict Entry Into the Tournament?

Actual performance could not account for the gender difference in choosing to perform in a tournament. However, beliefs about one's relative performance may be a major driving force for this gender effect, given that men form better beliefs about themselves than women.

We first examine whether beliefs are an important indicator for the participants' choice of incentive scheme. A probit analysis on entering the tournament as a function of one's guessed rank reveals that, for each gender, a participant with higher beliefs about their relative performance is significantly more likely to enter the tournament.[45] That is, for both, women and men, the more optimistic the beliefs about relative rank, the more likely they are to enter

[44] The marginal effects evaluated at a guess of 1, for a male with a tournament performance of 13 and a piece rate performance of 12, yields a significant female effect of -0.26 (s.e. 0.1, p=0.01), a significant coefficient on tournament performance 0.05 (s.e. 0.02, p=0.002) and a tournament-piece rate coefficient of 0.02 (s.e. 0.02, p=0.30).

[45] A probit regression on the probability to enter the tournament in the third task as a function of the guessed relative tournament performance in the second task reveals a coefficient of -0.24 (s.e. 0.11, p-value 0.035) for women and for men a coefficient of -0.21 (s.e.0.07, p-value 0.007), when the marginal effects are evaluated at a guessed rank of 1. We eliminated guesses of 4, and hence our sample is 39 men and 38 women.

the tournament. Figure 6 shows for each guessed rank the proportion of women and men that enter the tournament.[46]

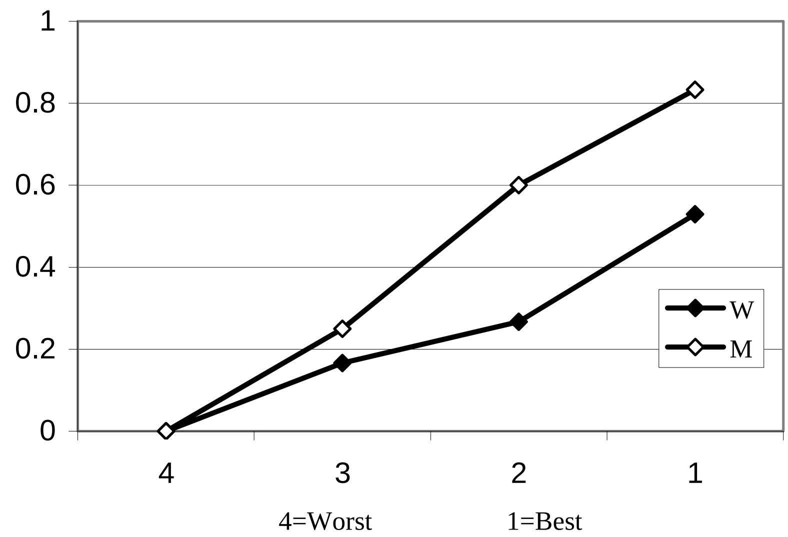

Figure 6: Proportion Entering the Tournament Conditional on Guessed Rank

A probit regression confirms that controlling for guessed rank (in the task 2 tournament) there remains a significant gender difference in choosing to enter the tournament in task 3. That is, even conditioning on beliefs about relative tournament performance, women enter the tournament less than men.

Table 9: Probit of Tournament Choice on Individual Characteristics

	Coefficient	Standard Error	p-value
Female	-0.28	0.11	0.009
Guessed Tournament Rank	-0.17	0.06	0.03

Notes to table: Dependent variable: task-3 compensation choice (1-tournament, 0-piece rate). Independent variables: a female dummy, guessed rank in task-2 tournament performance. The table presents marginal effects on the probability of selecting the tournament evaluated at a male guessing that he is ranked first. We eliminate guesses of 4 and hence are left with 39 men and 38 women.

Since actual performance could not account for the gender differences in choosing the tournament, we investigated the effects of the participants' beliefs about the relative ranking of their performance in the task 2 tournament. While women and men are both not very well

[46] Note that a participant with a point prediction of a guessed rank of 2 may still optimally choose to submit the piece rate result to a tournament payment scheme, for example, if the participant believes that she has a 40% chance to be best, and a 60% chance to be second.

calibrated in their beliefs about their guessed rank in their group of 4 players, men are significantly more overconfident than women. However, even conditioning on beliefs, there is a significant gender difference in choosing the tournament. That is while the difference in beliefs can account for some part of the gender difference in choosing the tournament, there remains a significant gender effect controlling for beliefs.

5.3.3: Decomposing the gender difference in tournament choice

We have seen that women and men differ in their believed tournament ranking, even controlling for performance. They also differ in the propensity to enter the tournament, for a given believed ranking. Next we assess the importance of each of these two effects in contributing to the gender difference in tournament entry as a function of past performance.

We have seen an overall gender effect in table 3 of 38%, that is a man with a performance of 13 in the tournament (and 12 in the piece rate) would have a 38% lower probability to enter the tournament if he were a woman. To assess the impact of belief formation, we estimate a probit regression of the decision to enter the tournament controlling for performance, the gender *and the guessed tournament rank*.[47]

With the control of the believed rank, the gender effect is 27%. That is, about 70% of the overall gender effect can be explained by women and men acting differently upon their beliefs, while 30% can be attributed to women and men forming different beliefs.

6. Why Women and Men may differ in their decision to enter a tournament

We found that men and women differ significantly in their propensity to enter a tournament, with about twice as many men choosing to compete compared to women. We saw that neither the tournament performance before the entry decision, nor the performance after the entry decision can account for this gender difference. However, it is not the absolute performance, but the relative performance in their group that determines the earnings from a tournament. Participants only know their absolute performance, and need to form beliefs about their relative performance. We elicited the participants' beliefs and found that men are

[47] The coefficient on tournament performance, and the increase in performance from the piece rate to the tournament remain small and not significant, when we evaluate the marginal effects at a man of a tournament performance of 13, a piece rate performance of 12 and a guessed tournament rank of 1. The marginal effect on the female dummy is -0.27 (s.e. 0.11, p=0.01) and on the guessed tournament rank it is -0.19 (s.e. 0.06, p=0.01). When we evaluate the marginal effects at a guess of 2, the marginal coefficient on the female dummy is slightly higher at -0.305.

significantly more overconfident than women. However, this gender difference in beliefs could not account for the gender difference in tournament entry. Even conditioning on beliefs women enter the tournament less than men.

We found that, compared to choices that maximize expected monetary earnings high performing women enter the tournament not often enough, that is, they seem to shy away from competition. On the other hand, low performing men enter the tournament too often, and seem to be unusually attracted to tournaments. In trying to understand what drives the gender gap in the choice of tournament, we will focus mostly on the high performing women, that is, we will try to ask why women shy away from competition. Many of the explanations, with a reverse argument, can be used to argue why men enter the tournament so much.

We will consider possible explanations that are of a general nature, that is, they have explanatory power for behavior in choosing a tournament, but are not specific to that decision, such as, for example, that women do not like uncertain payment schemes. We will also consider tournament specific explanations that arise form the fact that the performance is in a competitive environment, such as, for example, that women do not like tournaments.

Preferences

Women do not like to compete (Tournament Specific Explanation): Women may shy away from competition, simply because they do not like to be an environment in which they have to compete. Having a high cost for performing in a competitive environment may be a trait of women, which may not be at all correlated with their actual ability to perform in a competitive environment.[48] And, of course, alternatively it could be that men are simply drawn to competitive environments. This explanation is strictly about tournaments.

Risk Aversion (General Explanation): There is another aspect of preferences which will affect the decision to enter a tournament but is not just about tournaments. If women are more risk-averse than men, then a man and a woman with the same ability and performance (and the same beliefs about their relative performance) may make different choices, since tournaments involve uncertain payments. Whether there are gender differences in risk aversion, and how important they are, remains a subject of debate. Eckel and Grossman (2005) summarize the literature on gender differences in risk taking, and Eckel and Grossman (2002) find gender

[48] When we think of “psychic” costs of being in a tournament, then these costs would not affect tournament performance per se, simply the choice of performing in a tournament.

differences in choice of lotteries. The psychology literature is summarized in Byrnes, Miller and Shafer (1999). While the results are mixed, there seem to be somewhat higher degrees of risk aversion among women.

However, to explain the data from our experiment female participants would have to be extremely risk averse. For example, women (and men) with a task 2 tournament performance of 14 and higher have a probability of 47% and higher to win the tournament with such a performance. When we ignore the effort costs of playing in the tournament compared to a piece rate, the decision to enter the tournament becomes a gamble of receiving, per correct answer, either $2 with a probability of 47% or more, i.e. in expectation 94 cents (or more), or receiving 50 cents for sure. For participants who solve 14 correct answers that means a gamble of $ 28 with a probability of 47% (i.e. an expected value of $ 13), versus a sure gain of $ 7. Of the women who solve 14 or more, 8/12 do not enter the tournament (compared to 3/12 for men).[49] Similarly, men (and women) who solve 11 correct problems or less, have a probability of winning the tournament of 5.6% or less. Entering the tournament means receiving $ 2 per correct answer with a probability of 5.6% (or less), i.e. in expectation receiving 11 cents (or less) versus receiving 50 cents for sure. For participants who solve 11 correct answers this is a choice between receiving $22 with a probability of 5.6% (i.e. and expected value of $ 1.23) compared to receiving $5.5 for sure. 11/18 men who solve 11 or less take this gamble, compared to 5/17 women.[50] We are not aware of any paper that produces such extreme gender differences in risk aversion.

In general, if risk aversion were the main explanation, and costs of performing in a tournament do not vary widely, we would not expect men to enter the tournament with a higher probability than women for all performance levels, but rather the female entry decision to be a shift to the right of the male entry decision.

Feedback Aversion about one's relative performance (General Explanation): One other possible explanation for why women shy away from competition may be that women simply shy away from receiving feedback about their relative performance.

[49] This difference is marginally significant with a Fisher's exact test generating a two-sided p-value of 0.100.
[50] This difference is marginally with a Fisher's exact test generating a two-sided p-value of 0.092.

Forming point predictions of beliefs

We have seen that women and men form different beliefs about their relative tournament performance and this difference in beliefs can account for some portion of the gender difference in choosing the tournament.

Women have less optimistic views about their relative tournament performance (Tournament Specific Explanation): Women are less confident about their relative ability compared to men, when it comes to beliefs about their relative performance in a competitive environment, such as in the task 2 tournament. This could be because of a stereotype that women are not so competitive, or the fact that women may have been more stressed during the tournament, as would be suggested by the stereotype threat theory (Claude Steele 1997)[51]

Women always have less optimistic views about their relative performance (General Explanation): However, it could be that women have a lower perception of their relative ability, always, or always for this task, independent of the environment, i.e. the incentive scheme under which they performed.[52] (See e.g. Lichtenstein, Fischhoff and Phillips (1982), Beyer (1990) and see also Beyer and Bowden (1997)).

Certainty in and acting upon one's beliefs

However, neither actual tournament performance, nor beliefs about one's tournament performance can to a large extent explain why men and women differ in choosing the tournament incentive scheme, why women shy away from competition, and why men are so drawn to competitions.

One set of possible explanations why women and men act differently given their point prediction of guessed tournament rank, are the explanations we listed under the section headed preferences, namely: *Women do not like to compete (Tournament Specific Explanation)* and *Risk Aversion (General Explanation)*.

However, there are other reasons for the gender difference.

[51] Stereotype threat theory suggests that stereotyped individuals (e.g. women who are not supposed to be good competitors) who find themselves in a situation where they run the risk of confirming to the stereotype (i.e. in a tournament where they may lose) feel, next to other possible anxieties, this additional threat of confirming to the stereotype. This additional threat may harm performance; female participants may "choke" under this additional, female specific threat.

[52] It seems, however, that women are better calibrated than men in their beliefs about their relative tournament performance, so, the question could alternatively be posed as to why men have such high beliefs about themselves.

Uncertainty in beliefs about the relative tournament performance (Tournament Specific Explanation): It could be that women are less certain about their beliefs in their relative tournament performance than men. This could be for the same reasons with which women form lower point predictions about their relative ranking.

Another reason why women may act differently than men for a given point prediction of the believed task 2 tournament ranking, is that women think that their past tournament performance is not a good predictor for their future tournament performance. There is a rather extensive literature that asserts that women attribute past successes more to luck rather than inner attributes than men (and vice versa attribute past failures less to bad luck than men). (Beyer (1990) and Felder et al (1994)). In such a case, women and men would act differently conditioning on their belief about their point prediction of their relative ranking in the task 2 tournament.

Women are always less certain in beliefs about their relative performance (General Explanation). It could be that women are always less certain about their believed ranking. To test whether there is a basis for such beliefs, we estimate whether the performance of women is more volatile than the performance of men. We use the increase in performance between the piece rate and the tournament, as that is the only variance participants observed before making their choices. The average increase in performance is 1.45 for men (with a s.e. of 0.39) and 1.68 for women (with a s.e. of 0.36). And F-test on the equality of variance delivers p=0.64. There seems to be no evidence that the women's performance is more volatile than men's. Furthermore, even the change in performance from the tournament to the choice treatment has a similar pattern. The average increase in performance is 0.93 for men (with a s.e. of 0.42) and -0.15 for women (with a s.e. of 0.33). And F-test on the equality of variance delivers p=0.15. If anything, it is the change in male performance that exhibits higher variance.

Unreasonable Beliefs about future performance: It could be that men simply have unreasonable beliefs about future performance, especially when it is in a tournament environment. We have seen that both men and women have a higher performance in the tournament than in the piece rate, and participants (maybe especially men) may believe that they may continue to increase their performance. In this case we may expect that the decision to enter the tournament should be somewhat responsive to the increase in performance men and women have experienced between the piece rate and the tournament. However, as we have seen (cf table 3, and footnotes 15 and 17), the increase in performance has no correlation whatsoever with the decision to enter the tournament.

There are therefore a variety of possible explanations why women and men differ in the propensity by which they enter a tournament. With the last task, we want to examine two things. First we want to assess whether the general explanations, such as gender differences in risk attitudes and feedback aversion, a general relative lack of confidence of women, or a lack in the certainty of their beliefs are sufficient to by themselves generate a gender gap in the decision to choose a tournament compensation. Second, we want to test whether the general explanations can account for the gender differences in choosing to enter a tournament. That is, do any of the tournament specific explanations have a significant impact on the gender difference in entering a tournament, once we control for the general explanations?

One possible way to proceed would be to try to estimate the effects of all these general explanations separately, and then try to use our estimates as additional controls when we study the gender difference in choosing the tournament. However, it is well known that, for example, estimating parameters of risk aversion yields results that are very specific to the environment.[53] Furthermore, even if we would have good measures of risk aversion and of each of the other effects, we would still not know how they interact, for example it could be that women are more risk averse than men when deciding upon lotteries, but when it comes to making decisions which rely on one's beliefs, this could be a domain where women are even more risk averse.

Instead of measuring each effect separately, and making assumptions on how they interact and participants behave, we opt for a different strategy. We try to control for all these other effects at once, by having participants make a choice that is as close as possible to the choice in task 3, but which eliminates all tournament specific explanations. That is, we do not want to have women and men actually perform in a competitive environment after the choice. We also do not want their beliefs to be affected by a performance in a competitive environment, as women may simply form worse relative beliefs in such environments. Therefore, in our final treatment, we have women and men decide about the payment for a past performance.

[53] See for example Harbaugh, Krause and Vesterlund (2003).

7. Do Women and Men submit piece rate results to a Tournament?

7.1 Task 4: Submitting the Piece Rate to a Tournament Payment Scheme

In the fourth task, participants receive a payment on their task-1 piece rate performance. They can decide whether this payment is a piece rate payment (as in Task 1). Alternatively, they can decide to submit this performance to a competition. Participants win the competition, if they are the participant of the highest piece rate performance in their group independent on whether other participants decided to submit the piece rate performance as well.

As in the third task, where participants chose whether to enter the tournament or not, the decision on the payment scheme in task 4 does not affect any others participants payoff. We will use the decision to submit the piece rate to a tournament as a measure of the participants risk preference, willingness to act according to one's believed ranking etc.

This treatment allows us to measure whether in an environment where participants do not have to perform once more, and in which no tournament performance is involved, there still is a gender difference. Such a difference could be driven by feedback aversion, a lower confidence of women in their abilities, potential gender differences in risk aversion. We want to estimate, whether in such a situation, where the competitive aspect is almost eliminated (apart form the payment scheme), women still shy away from competition.

7.2 Do Women and Men Submit their Piece Rate Performance to the Tournament?

Before examining the participants choices, we estimate who should submit the piece rate to a tournament scheme, based on actual task-1 performance.

In the piece rate scheme, men and women do not have exactly the same probability of being the highest performer in a group of 2 men and 2 women drawn from a population with the distribution given by our experiment. Table 10 reports the probability of being the highest piece rate performer.[54]

Table 10: Probability of Winning Piece-Rate Tournament Conditional on Performance

	8	9	10	11	12	13	14	15	17	18	22
Women	1	3.6	11.4	21.6	33	49.4	66	81.4	--	93.9	--
Men	0.9	3.2	11.8	24.4	39.3	57.4	70.7	83.5	91.5	--	98.7

[54] For any given performance level, say 15 for a woman, we draw 10,000 groups consisting of 2 men and one other woman, where we use the sample of 40 men and women with replacement. We then calculate the frequency of wins. The exercise is repeated 100 times and we report the average of these win frequencies.

Overall, a man has a 29% chance to have the highest piece rate performance, while a woman has a 21% chance.[55] In our experiments, in the 20 groups there were 11 women and 11 men who were the highest performers in their group (obviously with some groups having ties in the highest performing participant).

With the tournament payment scheme, the winner (the highest performer in the group) receives $2 for every correct answer, while others receive no payment; in the piece rate payment, the participant receives 50 cents for every correct problem. Based on the piece rate performance, 30% of the women and 40% of the men (all those with a performance of 12 and higher) have higher expected earnings from submitting their performance to a tournament scheme as opposed to opting for a piece rate payment. Including participants who solve 11 problems – and are virtually indifferent between the 2 incentive schemes – the percentages are 45% for the men and 40% for the women.

In fact, 25% of the women decide to submit their piece rate performance, but 55% of the men, this difference is significant.[56]

We will now investigate whether the decision to submit the piece rate to a tournament payment scheme is driven by performance. We then investigate the formation of beliefs about relative rankings in the piece rate performance, and whether the beliefs can account for the gender difference in the choice of submitting the piece rate result to a tournament payment scheme.

7.3 Does Performance predict who submits the Piece Rate result to a Tournament?

To investigate if performance is a good predictor for the choice of the compensation scheme of the piece rate performance, Table 11 shows for women and men, the average piece

[55] This difference is not significant in a sample of 40 men and 40 women.

[56] The p-value of a Fisher's exact test is 0.012. Note that women and men have a lower propensity to submit the piece rate result to a tournament payment scheme than entering a tournament and then competing. This difference is however not significant for women (a Fisher's exact test yields p=0.465) and not for men (a Fisher's exact test yields p=0.162). One possible explanation for the albeit insignificant change is that in Task 4, only one of the four participants in a group can win the tournament, while in the decision in task 3, if all participants improve their performance by a lot, all participants could win. Another possibility is that participants have increased their performance after the piece rate scheme, and may not realize that this personal experience of an improved performance is a general experience that may be due to learning or the change in incentive scheme. There is a well known psychology literature that people attribute changes more to themselves than the environment: the fundamental attribution error (see e.g. Ross, L. (1977)). A failure to incorporate that others have experienced similar increases may lead to a relative under estimation of one's ability and hence to a lower decision to submit the piece rate result to a tournament (see also Moore and Small (2004)).

rate performance of participants who choose the piece rate payment scheme and of participants who choose to submit the piece rate performance to a tournament.

Table 11: Average Piece-Rate and Tournament Performance Conditional on Task-4 Choice

	Task-4 Choice	Piece Rate Performance	Difference
Women	Piece Rate	9.97 (0.54)	0.73
	Tournament	10.70 (0.76)	
Men	Piece Rate	9 (0.52)	3.05
	Tournament	12.05 (0.80)	
Men – Women	Piece Rate	-0.97	
	Tournament	1.75	

Notes to table: Difference shows the performance difference between participants who submitted their piece rate result to a tournament and those that did not.

For women there is no significant difference in performance between those who choose a tournament compensation scheme and those who choose a piece rate scheme.[57] A probit regression on the decision to submit the piece rate result to a tournament conditioning on the piece rate performance reveals that the coefficient on the piece rate performance is not significantly different from 0.[58] However, for men, the average performance of men who decide to submit their piece rate to the tournament is significantly higher than the average performance of those that do not.[59] A probit regression reveals that the piece rate performance has a significant correlation with the decision to submit the piece rate result to a tournament.[60]

Figure 7 shows the propensity of women and men to submit the piece rate to the tournament for each performance level of the piece rate, (we look at performance quartiles).

[57] A Mann Whitney test comparing the performance of women who submit their piece rate performance to the tournament and those that do not yields: for the piece rate: p=0.35 (t-test p=0.48).

[58] A probit regression on the probability to submit the piece rate result to a tournament as a function of the piece rate performance reveals a coefficient of 0.02 (s.e. 0.03, p-value 0.46) for the marginal effect evaluated at a performance of 11 for women.

[59] A Mann Whitney test comparing the performance of men who submit their piece rate performance to the tournament and those that do not yields: for the piece rate: p=0.01 (t-test p=0.004).

[60] A probit regression on the probability to submit the piece rate result to a tournament as a function of the piece rate performance reveals a coefficient of 0.08 (s.e. 0.03, p-value 0.01) for the marginal effect evaluated at a performance of 11 for men.

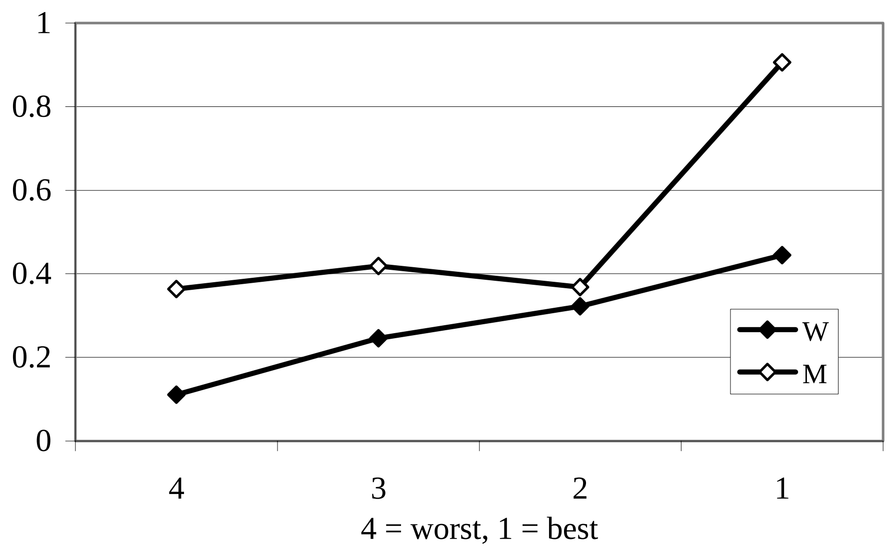

Figure 7: Proportion of participants that submit the Task 1 Piece Rate result to a tournament payment scheme conditional on the Task 1 performance quartile.

Of the participants who – in order to maximize expected earnings – should strictly submit the piece rate result to a tournament, that is who solve 12 of more, 14/16 (88%) of men submit the piece rate to a tournament, compared to 3/12 (25%) of women. A fisher's exact test confirms that this is a significant difference, p=0.001. Of the participants who solve 10 and less (and hence have a chance of less than 12% to win the tournament), 8/22 (36%) of men submit their piece rate to a tournament compared to 5/22 (23%) of women. This difference is not significant (p=0.33 with a Fisher's exact test).

The observation that conditional on the piece rate performance men submit the piece rate result more often to a tournament payment scheme than women is confirmed by a probit regression in table 12.

Table 12: Probit of Task-4 Compensation Conditional on Performance

	Coefficient	Standard Error	p-value
Female	-0.31	0.11	0.01
Piece Rate	0.06	0.02	0.01

Notes to table: Dependent variable: task-4 compensation choice (1-tournament and 0-piece rate). Independent variables: female dummy and piece rate performance in task 1. The table presents marginal effects on the probability of selecting the tournament evaluated at a male solving 11 problems in task-1

Women and men differ significantly in their propensity to submit their piece rate result to a tournament. This difference is driven by participants who have a piece rate performance of 12 and more (and have a probability to win the tournament of 33% and higher). The probit regression on the sample of participants whose performance is 10 or less yields a female dummy that is not significant, while on the sample of participants who solve 12 or more it is highly significant.[61]

We have seen that for men, but not for women, the piece rate performance does significantly predict the decision to submit the piece rate result to a tournament. There is a significant gender difference in submitting the piece rate result to a tournament, which is driven largely by the difference among high performing participants.

7.4: What are the effects (and differences) in beliefs about relative performance?

Before we assess the importance of the beliefs about the relative piece rate performance on the decision to submit the piece rate result to a tournament scheme, we study how women and men form those beliefs.

7.4.1 Formation of believed piece rate rank

To measure the effect of the piece rate performance on the guessed piece rate rank of participants, we use an ordered probit regression. We find that the women's performance is not significantly correlated with the formed beliefs about their relative ranking. For men, however, the correlation is significant.[62]

Before we study the actual distribution of guesses, we determine the distribution of perfectly calibrated individuals. Conditioning on gender only, having no prior about one's performance, optimal guesses are first and second for men, and second and third for women.[63] If participants know the distribution of performance for each gender, the resulting optimal guesses are reported in Table 13.

[61] A probit regression of the decision to submit the piece rate to a tournament on the piece rate performance and a female dummy yields, for participants who solve 10 or less in the piece rate, a coefficient on the piece rate of 0.03 (s.e. 0.05, p=0.6), and a female dummy of -0.17 (s.e. 0.14, p=0.23), when we evaluate the marginal effects at a male who solved 10 problems. For participants who solve 12 or more, the coefficients are, on the piece rate performance: 0.03 (s.e. 0.05, p=0.42) and a female dummy of -0.63 (s.e. 0.15, p=0.002), when evaluated at a male who solved 12 at the piece rate.

[62] An ordered probit regression on the guessed piece rate rank (for guesses 1, 2 and 3) as function of the piece rate performance reveals a coefficient of -0.09 (s.e. 0.07, p-value 0.21) for women and a coefficient of -0.33 (s.e.0.09, p-value 0.00) for men. For women, the average performance of participants who guess that they are ranked first is 10.9, for guesses of rank 2 it is 10.5 and for guesses of rank 3 it's 9.4. For men, the average performance is 12.75, 8.8 and 8.2 respectively. Optimal guesses should be 1 for a performance of 13 and higher, 2 for a performance of 11 and 12, and 3 for a performance of 9 and 10.

[63] Based on 10,000 artificially generated groups the likelihood of a woman being ranked first is 0.190, second 0.282, third 0.294, and last 0.233, the corresponding probabilities for a man for first is 0.262, second 0.253, third 0.250, and last 0.235. However, the expected loss from a wrong guess is not very high.

Table 13: Distribution of Actual and Optimal Guess of Piece-Rate Rank

	Men		Women	
	Actual Rank	Optimal Guess	Actual Rank	Optimal Guess
1: Best	11	12	11	6
2	12	6	9	10
3	9	11	11	15
4: Worst	8	11	9	9

Notes to table: Rank out of four. Actual rank is determined by the highest rank a participant could be placed in when controlling for ties.

Neither the actual ranks, nor the optimal guesses (which, for a given performance are the same for women and men) are significantly different across gender (Chi square tests yield p=0.88 and p=0.54 respectively). Table 14 shows the actual distribution of guesses.

Table 14: Distribution of Guessed Piece-Rate Rank

	Men		Women	
	Guessed Rank	Incorrect Guess	Guessed Rank	Incorrect Guess
1: Best	20	12	8	4
2	13	8	21	15
3	5	4	10	7
4: Worst	2	1	1	--
Total	40	25	40	27

Notes to table: Guessed rank out of four.

Men have more optimistic beliefs than women about their performance, a Chi-square test delivers p=0.029.[64] The distribution of incorrect guesses differs significantly across gender, a Chi square test yields p=0.047. Both, women and men are significantly overconfident. A Chi-square test that compares the guessed rank to the optimal guessed rank delivers p=0.005 for men and p=0.009 for women. The comparison to the actual rank (given by the particular groups of 4), delivers p=0.06 for men and 0.008 for women.[65]

To determine whether women and men form different beliefs conditional on their performance, we use an ordered probit to estimate the guessed piece rate rank as a function of the piece rate performance and a female dummy, using only guesses <4. The results in Table

[64] A Mann Whitney test comparing the beliefs of men and women yields a p-value of 0.02.
[65] This shows that participants seem not to be able to use characteristics of their specific group members, which would make the optimal guessed rank calculated by simulating on 10,000 artificially generated groups different from that optimal guessed rank based on the characteristics of that specific group.

15 reveal that women are significantly less optimistic about the ranking of their piece rate performance.

Table 15: Ordered probit of guessed ranks in the Piece rate Conditional on Performance

	Coefficient	Standard Error	p-value
Female	0.77	0.27	0.01
Piece Rate	-0.19	0.05	0.00

Notes to table: We consider only guessed piece rate ranks of 3 and lower, which leaves 30 men an 39 women.

Both women and men are overconfident in their beliefs, but men significantly more so than women.

7.4.2 Do beliefs predict who submits the piece rate to the tournament?

We have seen that the piece rate performance can predict which male participants but not which female participants submit their piece rate to a tournament payment scheme. The piece rate performance cannot account for the gender differences in that choice. It is the participants relative performance that determines the earnings from submitting the piece rate result. We have seen that men form significantly more optimistic beliefs about their relative performance than women. We now investigate whether the gender difference in submitting the piece rate can be explained by the fact that women form less optimistic beliefs than men.

For each gender beliefs are significantly correlated with the decision to submit the piece rate results.[66] Figure 8 shows, for each guessed piece rate rank, the proportion of women and men that submit their piece rate performance to a tournament payment scheme.

[66] A probit regression of submitting the piece rate to a tournament as a function of one's guessed rank (where we omit ranks of four), yields, a marginal effect of the believed rank (evaluated at a rank of 1) of -0.31 (s.e. 0.11, p=0.00) for men and -0.41 (s.e. 0.11, p=0.001) for women.

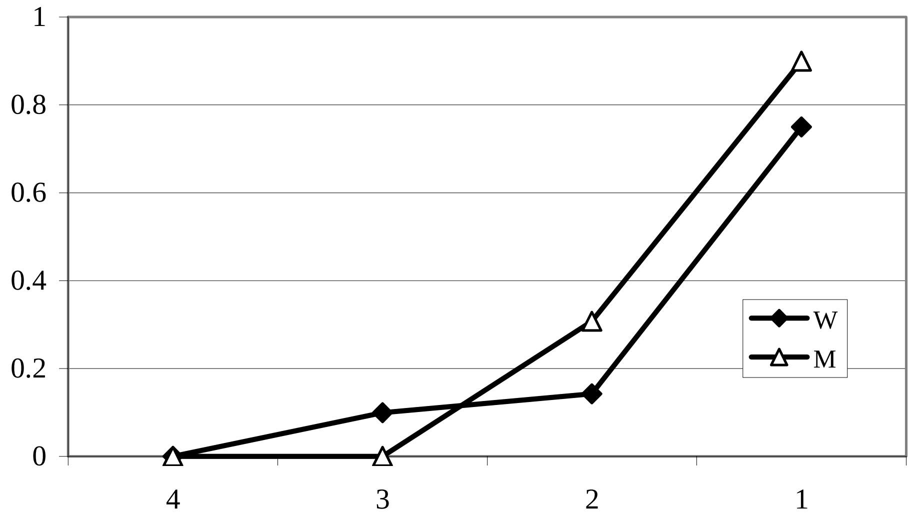

Figure 8: The proportion of participants that submit the piece rate result to a tournament payment scheme for every guessed piece rate rank

Women and men both have about a 60 percentage points higher probability to submit their piece rate result to a tournament payment scheme when they guess that they are the highest performer in their group, compared to when they are only the second highest performer in their group.

Furthermore the difference between the male and female decision is only about 15 percentage points. A probit analysis confirms that women and men do not differ significantly in their entry decision, once we control for the guessed rank. [67]

Table 16: Probit of Submitting the piece rate Choice on Individual Characteristics

	Coefficient	Standard Error	p-value
Female	-0.12	0.11	0.21
Guessed Piece Rate Rank	-0.32	0.08	0.00

Notes to table: Dependent variable: task-4 compensation choice (1-tournament, 0-piece rate). Independent variables: a female dummy, guessed rank in the task-1 piece rate. The table presents marginal effects on the probability of selecting the tournament evaluated at a male guessing that he is ranked first. We exclude guesses of 4, and are left with 39 women and 38 men.

[67] When we only consider guesses of 1 and 2, we have 62 participants, the marginal coefficient on the female dummy, evaluated at a guess of for males is: -0.14 (s.e.0.11, p=0.13) (with a significant coefficient on the guessed rank of -0.31 (s.e. 0.11, p=0.00)).

7.4.3 Decomposing the gender gap in the decision to submit the piece rate to a tournament

We have seen a gender gap in the decision to submit the piece rate result to a tournament, controlling for the piece rate performance (see Table 12). However, given a believed piece rate ranking, women and men did not differ significantly in their decision to submit the piece rate result. Now we want to assess the relative size of each effect.

When we restrict attention to participants with a believed piece rate rank of 3 or lower, the gender gap in the decision to submit the piece rate result controlling for the piece rate performance is 33%.[68] To assess the importance of the formed piece rate belief, we estimate a probit regression of the decision to submit the piece rate result as a function of gender, the piece rate performance, *and the guessed piece rate rank*.[69]

With the control of the believed piece rate rank the gender effect is 13%. That is about 40% of the overall gender effect can be attributed to men and women acting differently upon their beliefs, while 60% can be attributed to them forming different beliefs.

In task 4, we had women and men decide whether to submit their piece rate result to a tournament payment scheme or to a piece rate scheme. Actual piece rate performance is significantly correlated with the decision to submit the piece rate to a tournament payment scheme for men but not for women. Furthermore, controlling for the piece rate performance, there is a significant gender difference in submitting the piece rate, which is driven by a gender difference among high performance applicants. It is the participants' relative performance that affects the earnings of the decision about which participants can only form beliefs. The beliefs of men are correlated with performance, unlike the beliefs of women. While overall, both women and men are overconfident about their ability, men are significantly more optimistic than women. The overconfidence of men *is* sufficient to explain the gender difference in submitting the piece rate result to a tournament payment schedule.

[68] A probit regression on the decision to submit the piece rate result to a tournament as a function of gender and the piece rate performance for all participants with a believed piece rate rank of 3 or less yields marginal effects evaluated at a male who solves 12 that are: gender: -0.33 (s.e. 0.11, p=0.01), and piece rate: 0.05 (s.e. 0.02, p=0.02), for 38 men and 39 women.

[69] When we evaluate the marginal effects at a man of a piece rate performance of 12 and a guessed piece rate rank of 1, the marginal effect on the female dummy is -0.13 (s.e. 0.11, p=0.21), on the guessed piece rate rank is -0.31 (s.e. 0.09, p=0.00), and on the piece rate performance is 0.003 (s.e. 0.015, p=0.84).

8. Do women shy away from competition?

The decision to submit the piece rate performance to a tournament and the decision to enter a tournament share many aspects. In both cases the decision is about a piece rate payment scheme versus a tournament payment scheme. Furthermore, in both cases, after forming beliefs about one's relative performance, participants have to act according to these beliefs. Finally, a choice of tournament implies in both cases that participants receive feedback about their relative performance. The difference between the two decisions is that only when participants enter the tournament do they have to perform in a tournament, and only then do they have to assess and act upon their beliefs about their ranking in a tournament.

We have seen that there is a significant gender difference when participants decide whether to enter a tournament or not, and when participants form beliefs about their tournament ranking. We want to estimate whether gender differences in the decision to enter the tournament are driven solely by aspects that are present when participants decide whether to submit the piece rate, or whether there are additional gender differences when it comes to performing in a tournament and thinking about one's relative tournament performance.

The decision to submit the piece rate result to a tournament had a few differences to the decision to enter the tournament and then perform. While there was a significant gender difference in the rate with which participants submit the piece rate result for high performing participants, these differences were not significant among low performing participants. Furthermore, gender differences in belief formation about the piece rate performance were sufficient to eliminate gender differences in the decision to submit the piece rate result to a tournament, which was not the case when participants decided whether to enter a tournament.

This already suggests that the gender difference when participants decide whether to enter a tournament compared is different form the gender gap when women decide whether to submit a piece rate result. We now estimate the gender difference in entering a tournament using the decision to submit a piece rate result to a tournament as a control for aspects of the choice that are not solely tournament specific.

8.1: Do women only shy away from uncertain payments or feedback?

We want to test whether, controlling for performance and the decision whether participants choose to submit their piece rate performance to a competitive payment scheme, there still is a gender difference in the propensity with which participants choose to enter a competitive environment. We will use the participants' decision of whether or not to submit the piece rate performance to a tournament as a measure for their risk attitude and feedback aversion. With that measure we reevaluate the decision of participants to enter the tournament. A probit regression on the decision of whether or not to enter the tournament delivers:

Table 17: Probit of Task-4 Compensation Conditional on Performance

	Coefficient	Standard Error	p-value
Female	-0.25	0.11	0.01
Tournament	-0.00	0.01	0.98
Tournament – Piece Rate	0.02	0.02	0.16
Submitting the Piece Rate	0.34	0.12	0.004

Notes to table: Dependent variable: task-4compensation choice (1-tournament and 0-piece rate). Independent variables: female dummy and performance in task 2 and the increase form task 1 to task 2, and whether the participant submitted the piece rate. The table presents marginal effects on the probability of selecting the tournament evaluated at a male solving 13 problems in the task-2 tournament, 12 problems in the task 1 piece rate and having submitted the piece rate to the tournament.

That is, if using the decision of submitting the piece rate as a measure of gender differences that are not directly related to tournaments only, such as risk attitudes, receiving feedback on relative performance, we still find a significant (and large) gender effect. To address magnitudes of effects, we consider the marginal effect on the entry decision of a male with a tournament performance of 13 and a piece rate performance of 12 (who submitted the piece rate), where he to be a woman. The overall gender effect was 38% (see Table 3). Controlling for the decision to submit the piece rate, the gender effect is still 25%, that is only about 34% of the overall gender effect can be explained by general factors, such as risk attitudes, and the residual "competitive" component is 66%.[70]

Similarly, we consider the effect of controlling for the decision to submit the piece rate performance to a tournament, when we consider how the decision to enter a tournament correlates with the performance after the decision was made.

[70] When we restrict attention to participants who submitted the piece rate to a tournament, then a probit regression of the entry decision as a function of performance and a female dummy yields the following marginal effects when evaluated at a male who performs 13 in the tournament and 12 in the piece rate: female dummy: -0.29 (s.e. 0.18, p=0.096), tournament performance: -0.00 (s.e. 0.02, p=0.82), Tournament – piece rate: 0.01 (s.e. 0.03, p=0.67).

Table 18: Probit of Task-4 Compensation Conditional on Performance

	Coefficient	Standard Error	p-value
Female	-0.24	0.11	0.01
Task 3 Performance	0.003	0.01	0.76
Submitting the Piece Rate	0.29	0.11	0.01

Notes to table: Dependent variable: task-4compensation choice (1-tournament and 0-piece rate). Independent variables: female dummy and performance in task 3 and whether the participant submitted the piece rate. The table presents marginal effects on the probability of selecting the tournament evaluated at a male solving 13 problems in task-3 and having submitted the piece rate to the tournament.

Once again, the “competitive” component is 67% of the overall effect (see also Table 5).[71]

We therefore find that gender differences in risk attitudes or aversion to relative feedback performance cannot account for a vast portion of the gender difference in entering the tournament. Controlling for “noncompetitive” aspects, the marginal effect of gender on the decision to enter the tournament is still about 25%, when evaluated at the median performance of a male.

8.2: Gender differences in Belief formation

The participants expected outcome of the decision over incentive schemes are determined by their beliefs about their relative performance. We have seen that men form more optimistic beliefs than women when assessing their relative tournamnet performance. We now want to investigate whether, conditional on the believed piece rate rank, that is conditional on some measure of overconfidence, there is still a gender difference in how participants form beliefs about the relative tournament rank. That is, can the gender difference in tournament beliefs be eliminated when we use their formed piece rate beliefs as an independent measure of how participants form beliefs in general, for example as a measure of overconfidence.

[71] When we restrict attention to participants who submitted the piece rate to a tournament, then a probit regression of the entry decision as a function of task 3 performance and a female dummy yields the following marginal effects when evaluated at a male who performs 13: female dummy: -0.26 (s.e. 0.17, p=0.105), task 3 performance: -0.00 (s.e. 0.02, p=0.89).

Table 19: Ordered probit of believed tournament rank conditional on performance, gender, and the believed rank in the piece rate.

	Coefficient	Standard Error	p-value
Female	0.74	0.33	0.03
Tournament	-0.07	0.09	0.35
Tournament – Piece Rate	-0.25	0.09	0.00
Guessed Piece Rate Rank	0.82	0.28	0.00

Notes to table: Dependent variable: Believed tournament rank. Independent variables: female dummy and performance in the tournament and the increase in performance between the tournament and the piece rate, and the guessed rank in the piece rate. We omit the 6 participants who guessed a rank of 4 in either the tournament or the piece rate, that is we have 36 men and 36 women.

That is, even conditioning on the formed beliefs in the piece rate, there is an additional gender difference when participants form beliefs about their relative tournament ranking.

8.3: Do Women act differently upon their beliefs than Men?

We have seen that conditional on the participants believed ranking, there is still a significant gender gap in entering the tournament (of about 28%, when we evaluate the marginal effects of gender on a male with median beliefs, see Table 9). We want to estimate whether the gender difference is still significant and large once we control for the participants decision to submit the piece rate performance and their piece rate beliefs, that is once we control for "general" gender differences in deciding upon one's beliefs.

An ordered probit on the decision of whether or not to enter the tournament delivers:

Table 20: Probit of Task-4 Compensation Conditional on Believed ranking

	Coefficient	Standard Error	p-value
Female	-0.19	0.11	0.04
Guessed Tournament Rank	-0.16	0.07	0.00
Guessed Piece Rate Rank	0.13	0.08	0.04
Submitting the Piece Rate	0.47	0.17	0.002

Notes to table: Dependent variable: task-4 compensation choice (1-tournament and 0-piece rate). Independent variables: female dummy and the guessed rank in the task 2 tournament and the task 1 piece rate, and whether the participant submitted the piece rate. We eliminate all participants with one of the guesses being 4, which leaves us with 37 men and 37 women. The table presents marginal effects on the probability of selecting the tournament evaluated at a male with beliefs of 1 for both the tournament and the piece rate rank and having submitted the piece rate to the tournament.

The initial gender gap of 28% is reduced by about 32% to 19%, once we control for the decision to submit the piece rate. This provides some evidence, that the gender difference in deciding between entering the tournament or not seems not to be solely due to a general difference in how participants act upon their beliefs. [72]

9. Economic Consequences

We have seen significant gender differences in deciding whether to enter a tournament and whether to submit a piece rate performance to a tournament. We now estimate the expected costs of participants from deviating from money-maximizing choices.

To evaluate expected earnings from the decision in task 3 (i.e. whether to enter a tournament or not) we ignore performance costs (which we cannot measure). Furthermore, we assume that the performance is independent of the chosen incentive scheme. We consider two extreme ways of assessing the expected costs of entering a tournament. In columns 1 and 2 we use the task 2 tournament performance, which is the performance available to participants when they decide about entering the tournament. In columns 3 and 4 we use the actual task performance, assessing costs ex post, after participants performed.

While the magnitude of the costs is sensitive to the precise assumptions we make, the qualitative results are the same. The total costs of high performing participants who do not enter the tournament are higher for women than for men. On the other hand, the total costs of low performing participants who enter the tournament are higher for men than for women. By design the cost of a high performing participant who fails to enter the tournament is higher than the cost of a low performing participant who enters the tournament. We here higher total costs of women than of men.

The costs of men and women of submitting the piece rate to a tournament, or failing to do so follow a similar pattern. For high performers, women have larger total costs (of not submitting) then men, while among for performers total costs are higher for men than for women. Once more the total costs of women are higher than of men.

[72] When we condition the probit analysis only on participants who choose to submit the piece rate to the tournament, the marginal effects on the female dummy is somewhat larger,-0.32 (s.e.0.22, p=0.10), on the guessed tournament rank it is -0.10 (s.e.0.10, p=0.32) and on the guessed piece rate rank it is 0.22 (s.e. 0.17, p=0.14). This suggests that the effect of women and men acting differently upon beliefs in general has no explanatory power conditioning on participants who did submit their piece rate result to a tournament.

Table 21: Expected costs of decisions that do not maximize expected earnings

	Tournament entry: Costs based on Task 2 tournament		Tournament entry: Costs based on Task 3 performance		Submitting the piece rate: Costs based on Task 1 Piece Rate	
	Women	Men	Women	Men	Women	Men
Threshold	13	13	13	13	11	11
Should enter	12	12	9	20	12	16
Do not enter	8	3	6	4	9	2
Expected Cost	99.4	34.5	84.6	49.6	69.1	11.9
Average expected cost	12.4	11.5	14.1	16.5	7.7	5.9
Should not enter	24	22	24	19	24	22
Do enter	9	14	8	12	5	8
Expected Cost	32.9	56.5	28.9	43.8	15.0	28.1
Average expected cost	3.7	4.0	3.6	3.6	3.0	3.5
Total expected costs	132.3	91.0	113.5	93.3	84.1	40.0

Participants whose performance is at the threshold (and who are virtually indifferent between the two incentive schemes) are not included in the analysis. The costs of the decision in task 3 in columns 1 and 2 are based on the past task 2 tournament performance, while columns 3 and 4 use the task 3 choice performance. Columns 5 and 6 estimate the costs of decisions of submitting the piece rate to a tournament payment scheme.

10: Conclusions and Discussion

We conducted experiments to test the hypothesis that women shy away from competition. In our environment women and men perform equally well under both a piece rate incentive scheme and a competitive tournament scheme. Participants then decide in which incentive scheme they want to perform. We found that twice as many men as women enter the tournament. The choice of neither gender is strongly based on their actual performance (for both the performance before as well as the performance after the tournament entry decision). We found that men form more optimistic beliefs about their relative ranking (though both genders are overconfident), though this can only account for a small share of the gender gap in tournament entry.

Some of the possible explanations of this result are neither tournament specific, nor are they related to the fact that participants subsequently perform under a tournament compensation. For example, it could be that women are simply more risk averse, dislike receiving relative payoff information and generally are less optimistic about their relative performance (it could also be that women are less certain about their believed ranking). Other explanations for the gender differences in entry decisions are however tournament specific and rely on future performance in the tournament. For example, it could be that women dislike performing in a tournament, that they are less optimistic and precise in their beliefs on tournament performance, or that they expect to have low future performance.

In a last task we subject participants to a choice that excludes all aspects that relate to a future tournament performance, but retains all other characteristics. Specifically, we have participants decide whether to submit their past piece rate performance to a tournament or a piece rate scheme.

We find that even in this case, there is a significant gender gap in the decision to submit the piece rate result, and that this difference cannot be explained by the piece rate performance. Once again, men form significantly more optimistic beliefs about their ranking, though this time, when we control for the believed ranking, there is no significant gender difference in submitting the piece rate result to a tournament. That is, it appears that in this case, the gender difference in tournament choice is entirely driven by gender differences in beliefs about relative performance ranking.

Finally, in examining the tournament-entry decision we use the participants' decision of to submit the piece rate (jointly with the formed beliefs about their rank in the piece rate performance) as a control for aspects that are not related to a tournament or a future performance (e.g., general feedback aversion, potential gender differences in risk aversion, general overconfidence, etc.). We find that while these effects have some explanatory power, they do not account for the majority of the gender gap in entering the tournament.

That is, in a task where women and men perform equally well, we find that women shy away from competition, and that this behavior isn't simply caused by the payment scheme being uncertain, nor that they will receive feedback on relative performance. As much as women shy away from competition, men seem to be drawn towards them. This leads to lower earnings of women, especially high performing women, and monetary losses among low performing men (who enter the tournament too much).

The present paper is part of a research area that tries to understand why women are underrepresented in many high profile jobs and in whole professions. For example, women have a higher attrition rate from science and engineering, and it increases with academic rank.

Standard explanations include different preferences (or household or biological constraints) of women in terms of time to be invested in a job. An explanation for the lack of women in science and engineering is also possible differences in ability. An alternative explanation is discrimination, namely that the glass ceiling effect is man made, such that women may not be equally promoted and nurtured in science and engineering.

We studied an additional explanation, namely that women may be les “competitive,” less prone to select into competitions, but not because of differences in preferences over time invested in jobs, or differences in raw ability of performing in a task.

Gneezy, Niederle and Rustichini (2003) explored an environment in which there was no gender gap in performance in a piece rate scheme. However a mixed tournament created large gender differences in performance. In the mixed tournament, a few women were performing very highly, but a large number had a low performance, such that the bottom performance quintile was comprised of almost only women. Similar gender differences in competitive behavior have been found also by Larson (2005).

However, being able to perform well in a tournament, does not necessarily mean that a tournament would be preferred over a noncompetitive piece rate scheme. Indeed, in this paper we show that even when women perform as well as men in a competitive environment, women opt out of tournaments, while men opt in.

There is indeed evidence that, for example, the decision of women to quit sciences and engineering is not primarily due to ability. For example, a report entitled “Women’s Experiences in College Engineering” writes that “Many young women leave […] for reasons other than academic ability. These reasons can include their negatively interpreting grades that may actually be quite good, diminished selfconfidence, or reluctance to spend all of their waking hours ‘doing engineering.’” (Goodman, Cunningham and Lachapelle 2002). The report mentions that many women that left mentioned negative aspects of their schools’ climate such as competition, lack of support and discouraging faculty and peers. Similar effects have been found by Felder et al (1994).

It seems therefore that decisions of women to remain in male-dominated areas are not driven by actual ability only. In natural settings issues such as the amount of time devoted to the profession, and the desire of women to raise children may provide some explanations for the choices of women.

In this paper we examine an environment where women and men perform equally well, and where issues of discrimination, or time spent on the job do not have any explanatory power. Nonetheless we find large gender differences in the propensity to choose competitive environments. We feel that the effects we discover in the lab are strong and puzzling enough to call for a greater attention of standard economics to explanations of gender differences that so far have mostly been left in the hands of psychologists and sociologists.

References:

Altonji Joseph G., and Rebecca Blank "Race and Gender in the Labor Market" in O. Ashenfelter and D. Card, (eds), *Handbook of Labor Economics Volume 3c* Elsevier Science B.V. (1999): 3144-3259.

Andreoni, James, and Lise Vesterlund, "Which is the Fair Sex: Gender Differences in Altruism," *Quarterly Journal of Economics*, CXVI, (2001), 293 – 312.

Blau, Francine D., and Lawrence M. Kahn, 2004, "The US Gender Pay Gap in the 1990s: Slowing Convergence," NBER working paper 10853.

Bertrand, Marianne and Kevin F. Hallock, "The gender gap in top corporate jobs," *Industrial and Labor Relations Review*, LV (2001), 3-21.

Beyer, Sylvia, "Gender Differences in the Accuracy of Self-Evaluations of Performance," *Journal of Personality and Social Psychology*, LIX (1990), 960 – 970.

Beyer, Sylvia, and Edward M. Bowden, "Gender Differences in Self-Perceptions: Convergent Evidence From Three Measures of Accuracy and Bias," *Personality and Social Psychology Bulletin*, XXIII (1997), 157 - 172.

Byrnes, James P., David C. Miller, and William D. Schafer, "Gender Differences in Risk Taking: A Meta-Analysis," *Psychological Bulletin*, LXXV (1999), 367-383.

Costa-Gomes, Miguel, and Georg Weizsäcker, "Stated Beliefs and Play in Normal Form Games," working paper London School of Economics.

Della-Vigna, Stefano, Ulrike Malmendier, and Lise Vesterlund, "How robust is overconfidence: The effect of noisy and non-noisy signals," in preparation.

Eckel Catherine C. and and Philip J. Grossman, 2002, "Sex differences and statistical stereotyping in attitudes towards financial risk," *Evolution and Human Behavior* 23:281-295.

Eckel, Catherine C. and Philip J. Grossman, 2005, "Sex and Risk: Experimental Evidence", forthcoming, Handbook of Experimental Economics Results, Amsterdam, Elsevier Science, North-Holland.

Felder, Richard M., Gary N. Felder, Meredith Mauney, Charles E. Hamrin and E. Jacquelin Dietz, (1994), "A Longitudinal Study of Engineering Student Performance and Retention," Report NCSU-94A.

Fischbacher, U. (1999), z-Tree - Zurich Toolbox for Readymade Economic Experiments - Experimenter's Manual, mimeo

Ginther, Donna "Women in Economics: Moving Up or Falling Off the Academic Career Ladder?" *Journal of Economic Perspectives* (Summer 2004) 18:3, 193-214.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini, "Performance in Competitive Environments: Gender Differences," Quarterly Journal of Economics, August 2003, p. 1049-1074.

Harwarth. Irene, Mindi Maline and Elizabeth DeBra, "Women's Colleges in the United States: History, Issues, and Challenges", Office of Educational Research and Improvement at the U. S. Department of Education report, 1997.

Jan, Tracy; "A new gender divide: Coed school experiments with gender gap Coed Catholic school will try separating boys, girls in class," *Boston Globe*, April 9, 2005.

Jones, E. E. and Harris, V. A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology, 3,* 1-24

Krupka, Erin and Roberto Weber, 2004, "The Influence of Social Norms in Dictator Allocation Decisions," working paper.

Larson, Lia Christine, 2005, "When Girls Stop Competing Against Boys: An Experimental Analysis of the Competitive Behavior of Young Children", Harvard Honors Thesis.

Lichtenstein, Sarah, Baruch Fischhoff, and Lawrence Phillips, "Calibration and Probabilities: The State of the Art to 1980", in *Judgment Under Uncertainty: Heuristics and Biases*, Daniel Kahneman, Paul Slovic and Amos Tversky, eds., (Cambridge University Press: 1982)

Ledyard, John O., "Public Goods: A Survey of Experimental Research", in *Handbook of Experimental Economics*, John H. Kagel and Alvin E. Roth, eds. (Princeton, NJ: Princeton University Press, 1995)

Moore, D. A., & Small, D. A. (2004), "Error and bias in comparative social judgment: On being both better and worse than we think we are," Tepper Working Paper 2004-E1, Pittsburgh, PA.

National Science Foundation, Division of Science Resources Statistics, *Gender Differences in the Careers of Academic Scientists and Engineers*, NSF 04-323, Project Officer, Alan I. Rapoport (Arlington, VA, 2004).

Polachek, Solomon W., "Occupational Self-Selection: A Human Capital Approach to Sex Differences in Occupational Structure", *Review of Economics and Statistics*, LXIII (1981), 60-69.

Rosen, S. (1986) "The theory of equalizing differences", in O. Ashenfelter and R Layard, eds, Handbook of Labor Economics, Amsterdam, Elsevier Science, North-Holland.

Ross, L. (1977) The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (ed.), *Advances in experimental social psychology* (Volume 10, pp. 173-240), Orlando, FL: Academic Press.

Solnick Sara J., "Changes in Women's Majors from Entrance to Graduation at Women's and Coeducational Colleges," *Industrial and Labor Relations Review*, XLVIII (1995), 505-514.

Steele, Claude M., "A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance," *American Psychologist*, LII (1997), 613 – 629.

Appendix: Instructions

WELCOME

In the experiment today you will be asked to complete four different tasks. None of these will take more than 5 minutes. At the end of the experiment you will receive $7 for having completed the four tasks, in addition we will randomly select one of the tasks and pay you based on your performance in that task. Once you have completed the four tasks we determine which task counts for payment by drawing a number between 1 and 4. The method we use to determine your earnings varies across tasks. Before each task we will describe in detail how your payment is determined.

Your total earnings from the experiment are the sum of your payment for the randomly selected task, your $7-payment for completing the tasks, and a $5 show up fee. At the end of the experiment you will be asked to come to the side room where you will be paid in private.

Task 1 – Piece Rate

For Task 1 you will be asked to calculate the sum of five randomly chosen two-digit numbers. You will be given 5 minutes to calculate the correct sum of a series of these problems. You cannot use a calculator to determine this sum, however you are welcome to write the numbers down and make use of the provided scratch paper. You submit an answer by clicking the submit button with your mouse. When you enter an answer the computer will immediately tell you whether your answer is correct or not. Your answers to the problems are anonymous.

If Task 1 is the one randomly selected for payment, then you get 50 cents per problem you solve correctly in the 5 minutes. Your payment does not decrease if you provide an incorrect answer to a problem. We refer to this payment as the *piece rate* payment.

Please do not talk with one another for the duration of the experiment. If you have any questions, please raise your hand.

ARE THERE ANY QUESTIONS BEFORE WE BEGIN?

Task 2 - Tournament

As in Task 1 you will be given 5 minutes to calculate the correct sum of a series of five 2-digit numbers. However for this task your payment depends on your performance relative to that of a group of other participants. Each group consists of four people, the three other members of your group are located in the same row as you. The people immediately in front of you and behind you are in your group. If Task 2 is the one randomly selected for payment, then your earnings depend on the number of problems you solve compared to the three other people in your group. The individual who correctly solves the largest number of problems will receive $2 per correct problem, while the other participants receive no payment. We refer to this as the *tournament* payment. You will not be informed of how you did in the tournament until all four tasks have been completed. If there are ties the winner will be randomly determined.

Please do not talk with one another. If you have any questions, please raise your hand.

ARE THERE ANY QUESTIONS BEFORE WE BEGIN?

Task 3 - Choice

As in the previous two tasks you will be given 5 minutes to calculate the correct sum of a series of five 2-digit numbers. However you will now get to choose which of the two previous payment schemes you prefer to apply to your performance on the third task.

If Task 3 is the one randomly selected for payment, then your earnings for this task are determined as follows. If you choose the *piece rate* you receive 50 cents per problem you solve correctly. If you choose the *tournament* your performance will be evaluated relative to the performance of the other three participants of your group in the Task 2 -tournament. The Task 2-tournament is the one you just completed. If you correctly solve more problems than they did in Task 2, then you receive four times the payment from the piece rate, which is $2 per correct problem. You will receive no earnings for this task if you choose the tournament and do not solve more problems correctly now, than the others in your group did in the Task-2 tournament. You will not be informed of how you did in the tournament until all four tasks have been completed. If there are ties the winner will be randomly determined.

The next computer screen will ask you to choose whether you want the piece rate or the tournament applied to your performance. You will then be given 5 minutes to calculate the correct sum of a series of five randomly chosen two-digit numbers.

Please do not talk with one another. If you have any questions, please raise your hand.

ARE THERE ANY QUESTIONS BEFORE WE BEGIN?

Task 4 – Submit Piece Rate

You do not have to add any numbers for the fourth and final task of the experiment. Instead you may be paid one more time for the number of problems you solved in the Task 1 – Piece Rate. However, you now have to choose which payment scheme you want applied to the number of problems you solved. You can either choose to be paid according to the *piece rate*, or according to the *tournament.*

If the fourth task is the one selected for payment, then your earnings for this task are determined as follows. If you choose the *piece rate* you receive 50 cents per problem you solved in Task 1.

If you choose the *tournament* your performance will be evaluated relative to the performance of the other three participants of your group in the Task 1-piece rate. If you correctly solved more problems in Task 1 than they did then you receive four times the earnings of the piece rate, which is equivalent to $2 per correct problem. You will receive no earnings for this task if you choose the tournament and did not solve more problems correctly in Task 1 than the other members of your group.

The next computer screen will tell you how many problems you correctly solved in Task 1, and will ask you to choose whether you want the piece rate or the tournament applied to your performance.

Please do not talk with one another. If you have any questions, please raise your hand.

ARE THERE ANY QUESTIONS BEFORE WE BEGIN