

Do Women Shy Away from Competition?

Muriel Niederle and Lise Vesterlund

May 2005
Preliminary

Abstract

Competitive high ranking positions are largely occupied by men, and women remain scarce in engineering and sciences. Explanations for these occupational differences focus on discrimination and preferences for work hours and field of study. We explore an additional explanation, namely that women and men may differ in their selection into competitive environments. Men and women in a laboratory experiment perform a real task under a non-competitive piece rate and a competitive tournament scheme. Although there are no gender differences in performance under either of these compensations, there is a substantial gender difference when participants subsequently choose the scheme they want to apply to their next performance. Twice as many men as women choose the tournament over the piece rate. This gender gap in tournament entry is neither explained by performance before nor after the entry decision. Furthermore, while men are more optimistic about their relative performance, differences in beliefs only explain a small share of the gap in tournament entry. In a final task, we find that women are less likely to select tournament compensations even when they select it for past performance. In predicting tournament entry we use the compensation choice for past performance as a control for non-tournament specific gender differences (such as risk aversion, general feedback aversion and overconfidence), and we find a large residual gender effect.

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.¹ In academia, women are not well represented in math, science and engineering, and the difference increases with increases in academic rank. There also seems to be a leakage of women at the career ladder in economics.² These gender differences and their causes continue to be the subject of sometimes very heated debates.

Standard economic explanations for such occupational differences include preferences, ability and discrimination. Women may not select into competitive high profile jobs, or more technical fields because they do not enjoy the responsibilities associated with a managerial position, or because they dislike working in math, science or engineering. They may also avoid these jobs because they tend to have long work hours. Due to intrinsic preferences or intra-household bargaining women may choose to spend more time raising children. Secondly it has been suggested that the absence of women in some professions and higher ranks may be due to them having lower abilities or being less represented in the right tail of the ability distribution. While ability differences may account for the lack of women in certain professions and ranks, there is no consensus on the extent to which such differences translate into an occupational gender gap.³ Finally, a common explanation is discrimination or anticipated discrimination. Differential treatment of women and men with equal abilities may cause women to end up in occupations that differ from those of men.⁴

In this paper, we explore an additional explanation for the absence of women in higher ranking positions, namely that women may shy away from competitive environments. Such behavior not only reduces the number of women who enter tournaments, but also those who win tournaments. Hence it decreases the chances of women succeeding in competition for promotions and more lucrative jobs. This hypothesis resembles explanations that rely on pure preference differences between men and women. However it differs in that “perceived”

¹ For a recent review on the gender differences in wages in the 1990’s see Blau and Kahn (2004).

² See e.g., National Science Foundation (2004) and Ginther (2004). The absence of women in more technical fields extend far beyond academia, e.g., Hewitt and Seymour (1991) find that women only hold 10% of all jobs in physical sciences, engineering and math.

³ For a discussion on self-selection see Polachek (1981) and Rosen (1986)

⁴ Black and Strahan (2001), Goldin and Rouse (2000), see also Altonji and Blank (1999) and references therein.

preferences for competitive environments may be malleable, as they may depend on biased beliefs on ability, confidence, or feelings of competence.⁵

The psychology literature suggests that a lack of perceived competence and a fear of receiving negative signals about one's ability and performance may be an issue especially in areas where there exists a stereotype that questions one's ability in the domain.⁶ This would be the case for women who enter male-dominated domains, such as managerial positions, and fields such as math, science and engineering. Stereotype threat is one of the bases for support of single-sex schooling (see e.g., Jan (2005)).⁷ 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 will be willing to enter competitions against men and be successful at doing so.⁸

To test whether women shy away from competition, we use controlled laboratory experiments. These enable us to precisely measure performance, 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 sets of five two-digit numbers for five minutes. To determine the cause of possible gender differences in compensation choice (between a competitive and a non-competitive payment scheme) we opt for a task where we expect no gender differences in performance under either of the two schemes.⁹ Participants first perform the task under piece-rate compensation and then under a tournament. While participants are informed of their absolute performance after each task, they do not receive any feedback on their relative performance.

Having experienced both compensation forms participants then choose which of the two they want to apply to their performance of the next task, 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 problem. Specifically the participant's tournament-entry decision does not affect the payoffs of others, nor does it depend on the

⁵ Thus our line of research aligns well with the objective stated by the presidents at MIT, Princeton, and Stanford: "The question we must ask as a society is not 'Can women excel in math, science, and engineering?' -- Marie Curie exploded that myth a century ago -- but 'How can we encourage more women with exceptional abilities to pursue careers in these fields?'"

⁶ For the literature on stereotype threat theory and its implications see Steele (1997).

⁷ See Harwarth, Maline and DeBra (1997) and Solnick (1995) for benefits from single-sex schooling

⁸ Surprisingly there has been no 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) is not easily circumvented.

⁹ Gneezy, Niederle and Rustichini (2003) show that this may not be straightforward. We will tie our results to the past literature in the conclusion.

compensation choices she expects others to make. Despite there being no gender difference in performance under either compensation we find that twice as many men as women enter the tournament. Past performance cannot explain this gender difference and is generally a poor predictor of the participant's entry decision. The same holds for the participant's performance after the entry decision, once again performance is only weakly correlated with the entry decision, and it fails to account for the gender gap in tournament entry.

We also investigate if entry decisions may be driven by the participant's beliefs about her relative performance. We find that both genders are overconfident about their relative ranking, but that men are significantly more optimistic. While beliefs on relative rank help predict entry decisions, differences in believed ranking only account for a small share of the gender gap in tournament entry.

In addition to performance one can think of a series of potential explanations for why women may shy away from competitions. While some are specific to tournaments and performing under this compensation, others are not. Examples of non-tournament-specific explanations include that women may be more risk averse than men; that they in general may be more averse to receiving feedback on their relative performance; that they consistently may hold less optimistic beliefs than men; or, that they may not be very confident in the point predictions of their beliefs. We present the participants with one last task to determine whether these explanations by themselves generate a gender gap in entry decisions, or if tournament-specific explanations and future performance under this compensation is needed, i.e., to what extent is it the pressure of subsequently performing in a competitive environment that drives women away. 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 tournament-specific factors. 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 performances that occurred in a tournament setting. 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 either 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 original group of 4 people. As before women are significantly less likely to select the tournament compensation, 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. In contrast to the initial tournament-entry decision gender differences in beliefs can account for the gender gap in submitting the piece rate to the tournament. Controlling for individual beliefs one cannot reject that men and women are equally likely to submit the past piece-rate performance to a tournament. Thus gender differences in beliefs on relative rank explain why men are more likely to submit to the tournament.

We use behavior in this final task as a control for gender differences that may result from non-tournament-specific explanations, and we estimate the size of the residual gender difference in the decision to enter a tournament and subsequently perform. Including this control gender differences are still significant and large. That is, a sizable portion of the gender gap in tournament entry is driven by gender differences that are specific to the fact that participants after the compensation choice must perform under the selected compensation.

2. Experimental Design

We want to examine an environment where we can measure participants' performances under a competitive tournament and a non-competitive piece-rate scheme, and subsequently observe their choices between the two. This provides participants with experience of both compensation forms, and it enables us to determine the extent to which compensation choices are driven by ability. For our study we therefore conduct an experiment where participants solve a real task multiple times under varying compensation schemes. They first perform in a piece rate and then a tournament scheme. Participants are then asked to select which of these two schemes they want to apply to their next performance.

The task of our experiment is to add up sets of five 2-digit numbers. Participants are not allowed to use a calculator, but may write numbers down on scratch paper. The numbers are randomly drawn and each problem presented in the following way:

21	35	48	29	83	
----	----	----	----	----	--

Once the participant submits an answer on the computer, a new problem appears jointly with information on whether the former answer was correct.¹⁰ A record of the number of correct and wrong answers is kept on the screen. Participants have 5 minutes in which they may solve as many problems as they wish. We refer to this task as the 5-minute addition task.

The 5-minute addition task was selected because it requires both skill and effort, and because research suggests that there are no gender differences in ability on easy math tests.¹¹ The potential lack of a gender difference in ability is attractive as it enables us to better rule out ability differences as an explanation for gender differences in tournament entry.

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 randomly would be chosen for payment at the end of the experiment. By paying only for one task, we avoid that decisions in a given task may be used to hedge against outcomes in other tasks. Participants were informed of the nature of the tasks only immediately before performing the task. While participants know their absolute performance 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. Thus they do not know if their performance was better or worse than that of the other participants in their group. The specific tasks, compensations and order of tasks were as follows.

Task 1 – Piece Rate:

Participants are given the 5-minute addition task. If task 1 is selected for payment, participants receive 50 cents per correct answer.

¹⁰ The program was written using the software zTree (Fischbacher 1999).

¹¹ While males often score better on abstract math problems there is no gender difference in arithmetic or algebra performance, on computational problems women tend to score better than men (see Hyde, Fennema, and Lamon (1990) for a metaanalysis of 100 studies on gender differences in math performance).

Task 2 – Tournament:

Participants are given the 5-minute addition task. If task 2 is selected for payment, the participant who solves the largest number of problems in the group receives \$2 per correct answer, while the other participants receive no payment (in case of ties the winner is chosen randomly).

Note that the tournament is designed to secure that a participant with a 25% of winning the tournament receives the same expected payoff from the tournament as from the piece rate. This is important as we move onto the third task where participants once again are asked to perform the five-minute addition task, but this time select which of the two compensations they want to apply to their future performance.

Task 3 – Choice:

Before performing the 5-minute addition task participants select whether they want to be paid according to a piece rate, i.e., 50 cents for each correct answer, or a tournament. When the participant chooses tournament she received \$2 per correct answer if her score in task 3 exceeds that of the other group members in task 2, otherwise she receives no payment (in case of ties the winner is chosen randomly).

Winners of the task-3 tournament are determined based on the comparison relative to the other group members' task-2 rather than task-3 performance. One can think of this as competing against other participants who already performed.¹² This has several advantages; first, the performance of a player who enters the tournament is evaluated against the performance of participants who also performed under the tournament. Second, while beliefs regarding relative performance in a tournament may affect the decision to enter the tournament, beliefs regarding the choices of others will not. Thus we avoid an additional source of errors through biased beliefs about other participants' choices.¹³ A final benefit of this design is that a participant's choice does not affect the payment of any other participant,

¹² Many sports competitions are not performed simultaneously, e.g., downhill skiing.

¹³ For example, it could be that men believe that women will not enter the tournament, meaning that they face only one other competitor, greatly changing the odds of winning the competition.

thus there are no externalities associated with tournament entry.¹⁴ This allows us to exclude that women may shy away from competition because they by winning the tournament impose a negative externality on others.¹⁵ Effectively in task 3 participants face an individual decision problem which depends only on their ability to beat the task-2 performance of others and their preference for performing in a tournament.

If task 3 reveals that women are more reluctant to enter the tournament then we need to explain this difference. It may be that women simply do not like to perform in a competitive environment, or that they avoid 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. While some of these explanations are specific to the competitive tournament environment, others are not. We use a last task to determine whether there are still gender differences when we eliminate all aspects of the tournament choice that are specifically related to future performances in a competitive environment, while keeping all others. In the last task participants are subjected to a choice which is similar to that of task 3, but without using a tournament performance, and without having participants subsequently perform in a tournament.

Task 4 – Submit Piece Rate:

Participants do not have to perform in this task rather if selected for payment then the compensation depends on the number of correct answers in the task-1 piece rate. Participants choose which compensation they want to apply to their past piece-rate performance: a 50 cent piece rate or a tournament. They win the tournament and receive \$2 per correct answer if their task-1 piece-rate performance is highest of the participants in their group, 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. As in the

¹⁴ Our design allows for the possibility that there is no winner among participants who choose the tournament (if none of those entering the tournament beat the high score of their opponents). 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.

¹⁵ For a discussion on possible gender differences in altruism see e.g., Andreoni and Vesterlund (2001). See Ledyard (1995) for gender differences in social dilemma and public good games, as well as and Eckel and Grossman (2005) for a review of gender differences in experimental settings.

task-3 choice a participant's decision does not affect the earnings of any other participant, nor does it depend on the entry decisions of others. Thus task 4 is an individual decision task.

This final task allows us to see whether gender differences appear even when no future tournament performance is involved. Furthermore, we can use this task-4 decision as a control for the effect of uncertain payment, gender differences in assessing and acting upon one's beliefs in relative performance, and so on, these are all effects that are present in the task-3 choice, but are not unique to tournaments. With this decision as a control we estimate whether there is an additional gender difference when it comes to entering a competitive environment.

The participants' tournament-entry decisions (task 3 and 4) may be driven by absolute performance; however payments depend on relative performance. Since participants receive no feedback on relative performance during the experiment, entry decisions are likely affected by their beliefs about relative performance. To evaluate the effects of beliefs, and how participants form beliefs as a function of their performance and the payment scheme, we elicit beliefs on relative performance.

Belief-Assessment Questions:

At the end of the experiment participants are asked to guess their rank in the task-1 piece rate and the task-2 tournament. Participants have to pick one of the numbers 1,2,3 and 4, and are paid \$1 for each correct answer.¹⁶

At the end of the experiment, a random number from 1 to 4 is drawn to determine the task which is relevant for earnings. The experiment lasted about 45 minutes, and participants earned on average \$19.80.

3. No Gender Differences in Performance

The objective of this study is to examine whether, conditional on ability, women and men differ in their preference for performing under a piece-rate versus a tournament scheme. To eliminate ability differences as an explanation for potential gender differences in tournament

¹⁶ In case of ties in the actual 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.

entry, we want women and men to have similar performances under the two compensations. We first confirm that we found such a task.

Under piece-rate compensation there is no gender difference in performance. The average number of problems solved is 10.15 for women and 10.68 for men. Using a two-sided t-test this difference is not significant.¹⁷ Throughout the paper the reported test statistics refer to two-sided ttest, unless otherwise noted. We report only on Mann-Whitney tests if the conclusions reached from these differ from those of the t-test. Figure 1 shows the very similar cumulative distributions of the number of correct answers for women and for men.

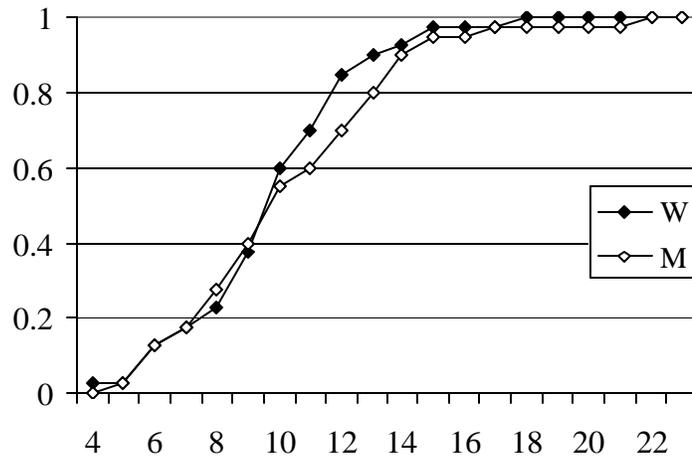


Figure 1: CDF of Piece-Rate Performance (Task 1)

The result is the same in the tournament. With the mean number of correct answers being 11.8 for women and 12.1 for men there is no significant gender difference in performance.¹⁸ The practically identical distributions of correct answers are shown in Figure 2.

¹⁷ The standard error is 0.44 for women and 0.55 for men. A two-sided t-test yields $p=0.459$. The average number of wrong answers is 2.8 for both men and women. Mean earnings are \$5.3 for men and \$5.08 for women.

¹⁸ The standard error is 0.48 for women and 0.43 for men. A two-sided t-test yields $p=0.643$. The average number of wrong answers is 3.0 for men and 2.3 for women this difference is significant ($p=0.033$). Mean earnings are \$6.85 for men and \$8.3 for women, this difference is not significant ($p=0.63$).

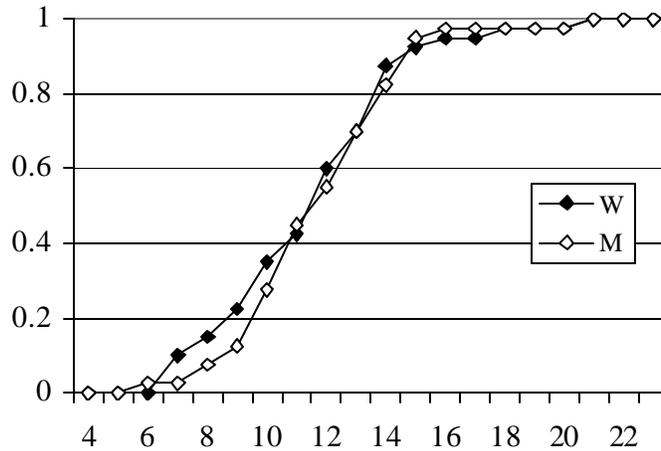


Figure 2: CDF of Tournament Performance (Task 2)

While the piece rate and tournament performances are highly correlated (with spearman rank correlations of 0.69 for women and 0.61 for men), the performance of both men and women is significantly higher under the tournament than the piece-rate.¹⁹ This improvement may be caused by learning or by the different performance incentives under the tournament.²⁰ The increase in performance varies substantially across participants, while this may simply be noise, it may also be due to some participants being more competitive than others. Note however that the increase in performance from the piece rate to the tournament does not differ by gender.²¹

The similar performance of men and women is reflected in their similar probabilities of winning the task-2 tournament. Of the 20 task-2 tournaments, 11 were won by women and 9 by men.²² To assess the probability of winning the tournament we randomly create four-person groups from the observed performance distributions. Conditioning only on gender the

¹⁹ The increase in number of correct answers 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.

²⁰ Malmendier, Della-Vigna, and Vesterlund (2005) have participants perform in series of 3 minute 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 initial learning may have some effect.

²¹ A t-test yields $p = 0.673$. In our sample, 37% of men and 25% of women experience no or negative change.

²² The 11 female winners had performances of 12, 13, 14, 14, 14, 14, 15, 15, 16, 18, 21, (average of 15.09). The performance of the 9 men is 12, 14, 14, 15, 15, 15, 15, 16, 21, (average of 15.22).

probability of winning the tournament is 26% for a man and 24% for a woman. As shown by Table 1 the probability of winning conditional on performance is also the same across gender.²³

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

Thus, our results suggest that we succeeded in selecting a task for which there is no gender difference in performance. After completing the first two tasks women and men have therefore had similar experiences and based on performance alone we would not expect a gender difference in the subsequent task-3 compensation choice.

4. Do Women Shy Away from Competition?

Having experienced both the 50-cent piece rate and the \$2 tournament participants are asked which of the two they want to apply to their task-3 performance. A participant who chooses the tournament wins the tournament if her number of correct answers in task 3 exceeds the number correct answers in task 2 by the other three members of her group. Thus choosing the tournament depends on beliefs regarding own ability and the other players' past tournament performance, but not on beliefs about the compensation choice of other participants.

Ignoring the costs of performing the task, the two compensations imply that a risk-neutral participant is indifferent between the two when her chance of winning the tournament is 25%. Thus from Table 1 we see that those with a performance of 14 and higher have higher expected earnings from the tournament. If the participant's task-3 performance is exactly like the task-2 performance this corresponds to 30% of the women and 30% of the men. 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, there are substantial difference in the compensation choice of women and men. While the majority of women prefer the piece rate, the majority of men prefer the tournament. With 35% of women

²³ 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.

and 73% of men selecting the tournament, there is a significant and substantial gender gap in tournament entry.²⁴ The remainder of the paper will focus on explaining this difference.

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 compensation choice in task 3.

5.1. Does Past Performance Predict Tournament Entry?

To investigate if past performance help predict the selected compensation we first compare the mean performance characteristics of participants who chose piece-rate to those who did not. Table 2 reports, by gender and the chosen compensation, the average number of problems solved under piece rate (task 1) and tournament (task 2), as well as the average increase in performance between the two.

Table 2: Performance Characteristics by Compensation Choice

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

Notes to table: standard errors in parenthesis

For women there is no significant difference in performance between those who do and do not enter the tournament.²⁵ A probit regression of tournament entry on past performance

²⁴ A chi square test yields a p-value of 0.001, and a fisher's exact test a p-value of 0.002.

²⁵ Using a t-test to compare the performance of the women who do and do not select the tournament generate the following test statistics: 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.

reveals that neither tournament performance nor the increase in performance from the piece rate to the tournament have a significant effect.²⁶

For men only the tournament performance is marginally higher for those who enter the tournament.²⁷ A probit regression of tournament entry 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, $p=0.15$).²⁸ That is, for men tournament performance has a very small and marginally significant effect on the entry decision. However, although higher tournament performing men are slightly more likely to enter the tournament, conditional on choice there is no gender difference in task-1 and task-2 performance.²⁹

Figure 3 shows the proportion of women and men who enter the tournament conditional on their performance quartile. Performance has, if any, only a small effect on tournament entry, and 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.

²⁶ A probit analysis of women's compensation choice on performance yields marginal effects on tournament performance of -0.01 (s.e. 0.03, $p=0.71$), on difference in performance between task 2 and 1 of 0.04 (s.e. 0.04, $p=0.31$) (evaluated at a tournament performance of 13 and a piece-rate performance of 12).

²⁷ 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.

²⁸ A probit of men's compensation choice on performance in task-2 and the difference between task 2 and 1 yields marginal effects 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 $p=0.83$), evaluated at a tournament performance of 13 and a piece-rate performance of 12.

²⁹ The p-values when comparing the performance of women to men who did enter the tournament, range from 0.28 to 0.48 when using t-tests. Similarly, when comparing women to men among participants who did not enter the tournament (p-values from 0.57 to 0.78).

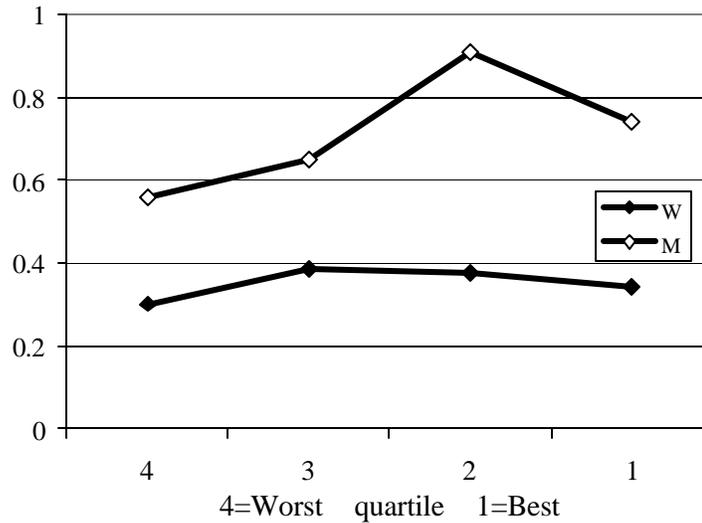


Figure 3: Proportion of Participants Entering Tournament Conditional on Task-2 Tournament Performance Quartile³⁰

Furthermore, among those whose past performance yield higher expected earnings in the tournament than piece rate (i.e., those solving 13 and more problems) significantly more men than women enter the tournament. Similarly men are significantly more likely to enter the tournament among those whose expected earnings are lowest in the tournament.³¹ The observation that, conditional on performance, men enter the tournament more than women is confirmed by a probit regression, reported in Table 3.

³⁰ 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.

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

Table 3: Probit of Tournament Choice on Individual Characteristics

	Coefficient	Standard Error	p-value
Female	-0.380	0.104	0.01
Tournament	0.015	0.018	0.41
Tournament – Piece Rate	0.015	0.023	0.50

Notes to table: Dependent variable: task-3 compensation choice (1-tournament and 0-piece rate). Tournament refers to task-2 performance, Tournament – Piece Rate to the change in performance between task-2 and task-1. The table presents marginal effects evaluated at a man solving 13 problems in the tournament and 12 in the piece rate.³²

Although both tournament performance and the improvement in performance have the expected positive signs, the effects are small and insignificant. While the decision to enter the tournament is not strongly related to the participants' performance under the two compensation schemes, the participant's gender is. Controlling for ability women are much less likely to select a competitive-compensation scheme.

One possible explanation for the observed gender difference in compensation choice may be that there is a gender difference in performance following the 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 task-3 performance conditional on the participant's compensation choice.

³² This evaluation point was selected because a risk-neutral individual solving 13 problems in the tournament is indifferent towards entering the tournament, and the average piece-rate performance was 12 for this group.

Table 4: Choice Performance Conditional on Compensation Choice

Compensation Choice		Average Performance	
		Choice (task 3)	Choice-Tournament
Women	Piece Rate	11.62 (0.62)	-0.15 (0.44)
	Tournament	11.79 (0.64)	-0.14 (0.50)
Men	Piece Rate	11.91 (0.91)	0.82 (0.44)
	Tournament	13.48 (0.83)	0.97 (0.55)

Notes to table: standard errors in parenthesis.

Conditional on gender Table 4 demonstrates no task-3 performance difference between those who do and do not enter the tournament.³³ Similarly, the participants who enter the tournament do not have a significantly different increase in performance in the choice task (task 3) relative to the tournament.³⁴ 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 who entered the tournament performed better than those who did not.³⁵

Although for a given gender, the task-3 performance does not predict which participants enter the tournament, it may help explain the gender gap in tournament entry. Figure 4 shows the cumulative distribution of the performance of women and men in task 3, independent of the chosen compensation scheme. While men solve an average of 13.05 problems in task 3, women solve only 11.68 problems. In contrast to our earlier results men solve significantly more problems than women in task 3.³⁶ 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.³⁷ This improvement in performance is significantly larger for men than it is for women.³⁸

³³ A t-test yields $p=0.86$ for women and $p=0.288$ for men.

³⁴ A t-test yields $p=0.99$ for women and $p=0.88$ for men.

³⁵ A probit analysis of the decision to enter as a function of the task-3 performance yields marginal effects of a participant that solves 13 in the task 3 task 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.

³⁶ The standard error in task-3 performance of men is 0.65 and of women 0.45. A ttest comparing the performance of women and men yields $p = 0.088$.

³⁷ p-values of a t-test for men and women are 0.03 and 0.65, respectively.

³⁸ A t-test comparing the increase in performance of women and men yields $p = 0.046$.

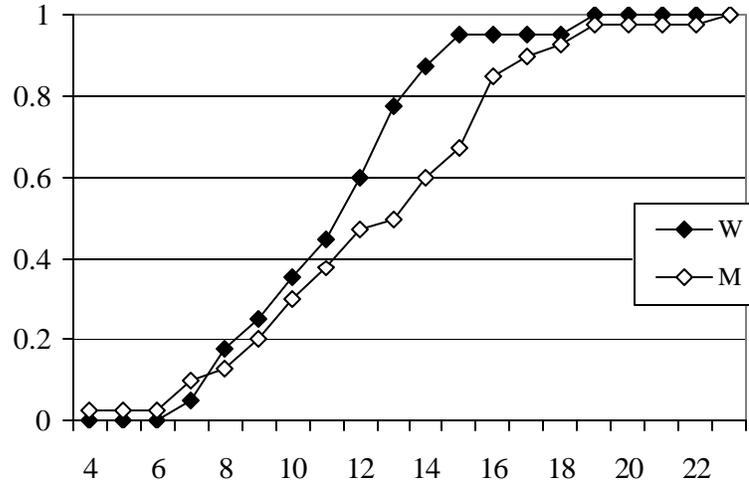


Figure 4: CDF of Choice Performance (Task 3).

The gender difference in task-3 performance however is an unlikely explanation for the gender gap in tournament entry. As seen in Figure 3 even men in the lowest performing quartile of the task-2 tournament (solving between 6 and 10 problems) have a higher propensity to enter the tournament than women in the top performing quartile (solving between 13 and 21 problems).³⁹ To assess the relationship between the task-3 performance with the compensation choice, i.e., to determine 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 for each performance quartile the proportion of participants that chose tournament. The gender difference in the tournament-entry decision conditional 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 enter the tournament. Whether we condition on those who have more or less than a 25% chance of winning the tournament, significantly more men than women enter the tournament.⁴⁰

³⁹ 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.

⁴⁰ Among participants who solve 13 and more (and have more than a 25% chance of winning the tournament), 17/21 (81%) of men and 6/16 (38%) of women enter the tournament, the difference is significant (a Chi-square test yields $p=0.01$). Similarly, among participants who have less than a 25% chance of winning, 12/19 (63%) of men and 8/24 (33%) of women enter the tournament ($p=0.05$).

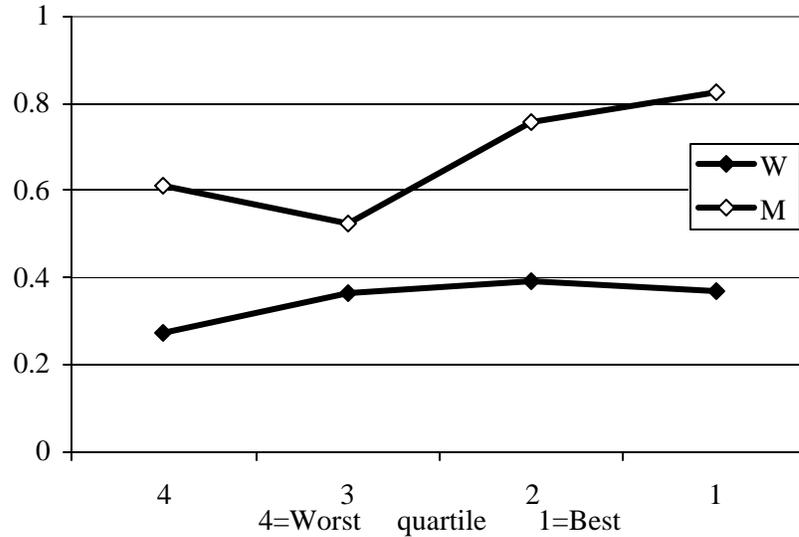


Figure 5: Proportion of Participants Entering the Tournament Conditional on Task-3 Performance Quartile.⁴¹

The probit regression in Table 5 confirms that while the participant’s gender significantly affects the tournament-entry decision, the future task-3 performance does not. That is, the gender difference in task-3 performance cannot explain the gender gap in tournament entry. Controlling for the unique increase in male performance in the third task women are significantly less likely to enter the tournament.

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

	Coefficient	Standard Error	p-value
Female	-0.357	0.106	0.002
Task-3 performance	0.015	0.014	0.311

Notes to table: Dependent variable: task-3 compensation choice (1-tournament and 0-piece rate), the table presents marginal effects evaluated at a man solving 13 problems in task-3.

5.3 Is the Entry Decision Determined by Beliefs?

Actual performance cannot explain the gender gap in compensation choice. This may be because tournament entry depends not on an individual’s absolute ability, but on how the ability ranks relative to the other group members. While participants are aware of their absolute

⁴¹ For each performance quartile we report the propensity by which men and women in this quartile enter the tournament.

performance, they do not know their relative ranking. Hence the compensation choice may depend on her belief about relative performance. We elicited these beliefs at the end of the experiment by asking participants 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.⁴²

We start by examining if men and women of equal ability differ in their assessment of their relative performance. We then ask whether these potential confidence differences can account for the gender difference in tournament entry. That is conditional on beliefs about relative performance are men and women equally likely to choose the tournament.

5.3.1. Do Women and Men Form the Same Beliefs?

Let us first characterize the distribution of guessed ranks we should expect to see if participants were perfectly calibrated. Due to the tie-breaking rule participants, who only know they are drawn from the performance distribution, have an incentive to guess that they are ranked second or third.⁴³ 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.⁴⁴

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 can be correct in the case of ties.

⁴² For example, in the event of a two-way tie for first rank, a guess of either first or second rank warrants the \$1 payment. 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. The elicited beliefs indicate that this was not a motive for the majority of participants.

⁴³ 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.

⁴⁴ 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).

These distributions differ substantially from the relative ranking participants believed they had. Table 7 shows that even though the tournament performance of men and women is basically identical, men are more optimistic about their relative performance: 75% of the men think they are best in their group of 4, compared to 43% of the women. The guesses of women and men differ significantly from one another.⁴⁵ Table 7 also reports the distribution of wrong guesses, and shows that men are more likely to incorrectly guess that they are ranked first. The distributions of incorrect guesses differ significantly across gender.⁴⁶ 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.⁴⁷

Table 7: Distribution of Gessed Tournament Rank

	Men		Women	
	Gessed Rank	Incorrect Guess	Gessed 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: Gessed rank out of four.

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 gessed rank as a function of tournament performance, the increase in performance and a female dummy. The results in Table 8 reveal that while participants with high tournament performance think they have higher relative performance, women are significantly less optimistic about the ranking than men.⁴⁸

⁴⁵ A Chi-square test delivers $p=0.025$. A t-test comparing the mean beliefs of men and women yields $p=0.024$.

⁴⁶ 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.

⁴⁷ A Chi-square test of independence between the distribution of gessed rank and optimal gessed rank yields $p=0.0008$ for men and $p=0.102$ for women, the test of independence between distribution of gessed rank and actual rank yields $p=0.0001$ for men and $p=0.057$ for women.

⁴⁸ The marginal effects evaluated at a guess of 1, for a man with a tournament performance of 13 and a piece-rate performance of 12, yields -0.26 on female (s.e. 0.1, $p=0.01$), 0.05 on tournament performance (s.e. 0.02, $p=0.002$) and 0.02 on tournament-piece rate (s.e. 0.02, $p=0.30$).

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 for guesses of ranks 1, 2, and 3.⁴⁹

Both women and men base their beliefs weakly on performance.⁵⁰ However, conditional on performance, men are significantly more optimistic about their relative performance.

5.3.2. Do Beliefs Predict Entry Into the Tournament?

While actual performance does not account for the gender gap in tournament entry, the substantial male overconfidence may help explain the difference. We first examine if beliefs are an important indicator for the participants' compensation choice. A probit analysis of tournament entry as a function of one's guessed rank reveals that, for each gender, participants with higher beliefs about relative performance are significantly more likely to enter the tournament.⁵¹ Figure 6 shows for each guessed rank the proportion of women and men that enter the tournament, and confirms that, independent of gender, the more optimistic the beliefs about relative rank, the more likely is tournament entry.⁵²

⁴⁹ 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.

⁵⁰ This result is confirmed in separate probit regressions for men and women. The average performance of a man 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.

⁵¹ A probit regression on the probability of entering the tournament in the third task as a function of the guessed relative tournament performance in the second task reveals a marginal effect of -0.24 (s.e. 0.11, p-value 0.035) for women and for men an effect of -0.21 (s.e.0.07, p-value 0.007), when evaluated at a guessed rank of 1. We eliminate guesses of 4, and hence our sample is 38 women and 39 men.

⁵² 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.

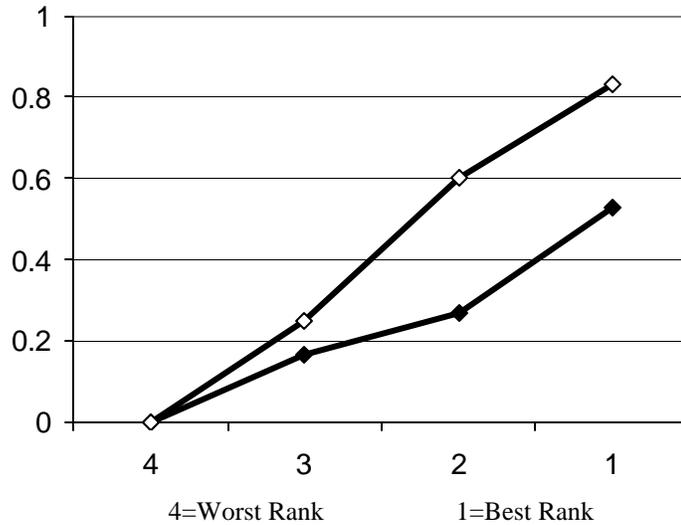


Figure 6: Proportion Entering the Tournament Conditional on Gussed Rank

Figure 6 also demonstrates that conditional on believed ranking women are less likely to enter the tournament. A probit regression confirms women remain significantly less likely to enter the tournament when we in addition to performance also control for gussed rank (in the task-2 tournament)

Table 9: Probit of Tournament Choice on Individual Characteristics

	Coefficient	Standard Error	p-value
Female	-0.278	0.113	0.012
Tournament	-0.002	0.016	0.897
Tournament – Piece Rate	-0.001	0.020	0.940
Gussed Tournament Rank	-0.181	0.064	0.005

Notes to table: Dependent variable: task-3 compensation choice (1-tournament and 0-piece rate). The table presents marginal effects evaluated at a man who thinks he is ranked first, and who solves 13 problems in the tournament and 12 in the piece rate. Guesses of 4 are eliminated resulting in a sample of 38 women and 39 men.

While neither women nor men are well calibrated in their beliefs about their gussed rank in their group of 4 players, men are significantly more overconfident than women. Nonetheless a significant gender difference in tournament entry remains when controlling for beliefs. 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.

In Table 3 we saw that an overall gender effect in table 3 of 38% on tournament entry, that is a man with a performance of 13 in the tournament (and 12 in the piece rate) would have

a 38% lower probability of entering the tournament if he were a woman.⁵³ Table 9 above shows that including the control for guessed tournament 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. Alternative Explanations for the Gender Gap in Tournament Entry.

Our results thus far show that neither performance before nor after the entry decision can explain the substantial gap in tournament entry. We also find that the substantial male overconfidence only explains a small share of the difference. From a payoff-maximizing perspective high performing women enter the tournament too rarely, and low performing men enter the tournament too often.

In this section we consider a number of alternative explanations for why women shy away from competition, while men embrace it. In examining these we distinguish between reasons that are and are not unique to the tournament compensation. For example, a non-tournament specific explanation may be that women inherently dislike uncertain payment schemes, while a tournament-specific explanation may be that women do not like the pressure of performing in a competitive environment. We consider in turn explanations that are related to preferences, confidence in ability and precision of ones beliefs.

Preferences

Women do not like to compete (tournament specific): Women may shy away from competition, simply because they dislike being in an environment where they have to compete. The pressure of future competition may impose a psychic cost and deter women from tournaments. Of course the reverse argument is that men may get a psychic benefit from performing in competitive environments and therefore be drawn to them. This explanation is unique to tournament compensation. Note that the discomfort or comfort of competition need not be correlated with the ability to perform in a competitive environment.⁵⁴

⁵³ The coefficient on female is also -0.379 when we exclude those with guesses of 4.

⁵⁴ While “psychic” costs and benefits of a tournament may deter entry, it need not affect tournament performance.

Risk aversion (general): As tournaments involve uncertain payoffs, risk attitudes are another aspect of preferences which may affect the compensation choice. Note that uncertain payoffs are not unique to the competitive tournament environment. If women are more risk-averse than men, then a man and a woman with the same ability, performance and beliefs about relative performance may prefer different compensations.⁵⁵

In considering this explanation it is important to note that the gender difference in risk aversion would have to be very large to explain the observed gap in tournament entry. For example, for participants who have 14 or more correct answers the chance of winning the tournament is 47% and higher. Presuming that one can maintain the performance in task 3 and ignoring the potential costs of performing in the tournament, the decision to enter the tournament becomes a gamble of receiving, per correct answer, either \$2 with a probability of 47% (or more), or receiving 50 cents for sure. For participants who have 14 correct answers that means a gamble of a 47% chance of \$28 (i.e., an expected value of \$13), versus a sure gain of \$7. Of the participants who solve 14 problems or more, 8/12 of the women and 3/12 of the men take this gamble.⁵⁶ Similarly, for participants who have 11 or fewer correct answers the chance of winning the tournament is 5.6% or less. Thus entering the tournament means receiving \$2 per correct answer with a probability of 5.6% (or less) versus receiving 50 cents for sure. For participants who solve 11 correct answers this is a choice between a 5.6% chance of winning \$22 (i.e., an expected value of \$1.23) compared to receiving \$5.5 for sure. Of the men who solve 11 problems or less 11/18 take this gamble while only 5/17 women do.⁵⁷ We are not aware of any studies that find such extreme gender differences in risk aversion.

Furthermore if risk aversion is the main explanation for the gender gap in tournament entry then we would not expect men to enter the tournament with a higher probability than women for all performance levels, but rather that the female-entry decision was shifted to the right of the male-entry decision.

Feedback aversion about one's relative performance (general and tournament specific): Another possible explanation for why women shy away from competition may be that they

⁵⁵ While there is some debate on whether there is a gender difference in risk attitudes, most studies find either no gender difference or that women are more averse to risk than men. Eckel and Grossman (2005) summarize the literature on gender differences in risk taking, and Eckel and Grossman (2002) find gender differences in choice of lotteries. The psychology literature is summarized in Byrnes, Miller and Shafer (1999).

⁵⁶ This difference is marginally significant with a two-sided Fisher's exact test ($p = 0.100$).

⁵⁷ This difference is marginally significant with a two-sided Fisher's exact test ($p = 0.092$).

dislike receiving feedback about their relative performance. While information on relative performance is not unique to the tournament, it may be that the participants are more averse to receiving feedback on relative performance in a competitive setting. Thus some component of feedback aversion may be unique to tournaments.

Point predictions of beliefs

Women are less optimistic about their relative performance (general and tournament specific):

Our results show that women are less optimistic about their relative tournament performance and that this difference can account for some portion of the gender gap in tournament entry. However a lower level of female confidence need not be unique to a competitive environment, and may be present independent of the incentive scheme (see e.g., Lichtenstein, Fischhoff and Phillips (1982), Beyer (1990) and Beyer and Bowden (1997)).⁵⁸ On the other hand it is entirely plausible that women are even less confident in competitive environments, such as the task-2 tournament. This could be because of a stereotype that women are not so competitive, or that women may be more stressed during the tournament (Steele 1997).⁵⁹ Thus some component of lower female confidence could be specific to the tournament.

Certainty in and acting upon one's beliefs

Uncertainty in beliefs about relative performance (general and tournament specific): In addition to being less optimistic about their relative ranking, it may be that women also are less certain about how correct their believed ranking is. That is, independent of the environment women may be more hesitant in responding to their beliefs because they perceive these as being imprecise. As a test of the basis of this belief we determine whether the performance of women is more variable than that of men. We examine the increase in performance between the piece rate and the tournament, as that is the only variance participants observe before making their compensation 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). The increase in performance does not differ across gender and an F-test on the equality of variance delivers $p=0.64$. Thus there is no

⁵⁸ It seems, however, that women are better calibrated than men in their beliefs about their relative tournament performance, hence the more appropriate question may be why men think so highly of themselves.

⁵⁹ Stereotype threat theory suggests that stereotyped individuals (e.g., women who are supposed to be poor competitors) who find themselves in a situation where they run the risk of confirming the stereotype (i.e., in a tournament where they may lose) may feel additional performance anxiety for fear of confirming the stereotype. This additional threat may harm female performance as they may “choke” under the pressure.

evidence that women receive more volatile performance information prior to their choice. Furthermore, the change in performance from the tournament to the choice task 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). Men improve their performance significantly more than women, and an F-test on the equality of variance delivers $p=0.15$. If anything, male performance is more variable.

While uncertainty about an individual's beliefs does not only affect tournament entry, such uncertainty may be more severe in a competitive setting where women are perceived to do poorly. Thus there may be a tournament-specific component to uncertainty in beliefs.

Biased beliefs about future performance (tournament specific): Another reason why women may act differently than men for a given point prediction of the believed task-2 ranking, is that women may think that their past tournament performance is a poor predictor for future tournament performance. A rather extensive literature asserts that women are more prone to attribute past successes to luck than to inner attributes (and past failures less to bad luck).⁶⁰ If this is the case then women and men will act differently conditional on their believed ranking in the task-2 tournament.

Another reason for biased beliefs about future performance in a tournament environment may be that men overestimate how well they will do in future tournaments. With participants having a higher performance in the tournament than in the piece rate, men may be more prone to believe that they will continue to increase their performance. Note however that if this were the case we would expect the tournament-entry decision to be correlated with the increase in performance participants experience between the piece rate and the tournament. However, as seen in Table 3 (and footnotes 28 and 30), the increase in performance is not correlated with tournament entry.

There are a variety of possible explanations for why women and men differ in the propensity by which they enter a tournament. Ideally we would like to determine the extent to which the tournament environment alone pushes women away from competition. One possible way of proceeding is to separately estimate the effect of each of these general explanations, and then use our estimates as additional controls when we study the gender difference in

⁶⁰ Beyer (1990) and Felder et al (1994).

tournament entry. However, it is well known that, for example risk attitudes are sensitive to the environment in which they are elicited.⁶¹ Furthermore, even if we had good measures of risk aversion and of the other effects, we would not know how they interact. For example it may be that women are more risk averse than men when choosing between lotteries, but that they appear even more risk averse when making decisions that rely on one's beliefs.

Instead of measuring each effect separately and making assumptions on how these interact we opt for a different strategy. We try to simultaneously control for all the general effects, 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. Specifically, in our final task, we have participants select compensation for their past piece-rate performance (task 1). Thus the choice is not based on performance in a competitive environment, and it does not require that participants perform after their decision.

7. Submitting Piece-Rate Results to a Tournament

If the fourth task is selected for payment the participant is paid for her task-1 piece-rate performance, however this time she decides whether she wants the 50-cent piece rate or the \$2 tournament compensation. If the tournament is chosen, the piece-rate performance is submitted to a competition among the piece-rate performances of the other participants in the group (independent of their compensation choice). A tournament is won if an individual's performance exceeds that of the other three players.

Task 4 enables us to first assess whether general explanations, such as gender differences in risk attitudes, feedback aversion, confidence, and certainty of beliefs by themselves are sufficient to generate a gender gap in compensation choices. Second, we can use the task-4 decision to test whether the tournament-specific factors cause a gender difference in tournament entry (task 3), or if general explanations can account for the entire gender gap. That is, does a gender difference in tournament entry remain once we control for the general explanations?

⁶¹ See for example Harbaugh, Krause and Vesterlund (2003).

7.1 Do Women and Men Submit their Piece-rate Performance to the Tournament?

Before examining the participants' choices, we use the task-1 performance to determine who should submit the piece rate to a tournament scheme. In the piece rate men and women have similar but not exactly the same probability of being the highest performer in a randomly drawn group of 2 men and 2 women. Overall the chance of getting the highest piece-rate performance is 29% for a man and 21% for a woman.⁶² In our 20 groups 11 women and 11 men were the highest performers in their group (incl. two cases of ties). Table 10 reports the probability of being the highest piece-rate performer and winning the tournament.⁶³

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

With the per problem compensation of 50 cents under the piece rate and \$2 to the tournament winner, those with a performance of more than 11 have higher expected earnings from submitting to a tournament, corresponding to 30% of the women and 40% of the men. Including participants who solve 11 problems – and are virtually indifferent between the two compensation schemes – the percentages are 40% for the women and 45% for the men.

The actual choices differ from these as 25% of the women and 55% of the men decide to submit their piece-rate performance to the tournament. This difference is significant.⁶⁴

In explaining this difference we first see if the compensation choice is driven by performance, and then whether it is driven by beliefs on relative piece-rate performance.

⁶² This difference is not significant in a sample of 40 men and 40 women.

⁶³ 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.

⁶⁴ The p-value of a Fisher's exact test is 0.012. Note that participants are more reluctant to submit the piece rate result to a tournament than they were to enter a tournament and then competing. This difference is neither significant for women (a Fisher's exact test yields $p=0.465$) nor for men (a Fisher's exact test yields $p=0.162$). One possible explanation for the albeit insignificant change is that in Task 4, there can only be one winner in each group of four, while in task 3 all participants can win the tournament, provided they improve their performance by a lot. Another possibility is that participants who increased their performance after the piece-rate scheme, may not realize that this personal experience is common for all participants. From psychology we know 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 under estimation of one's ability and hence to a lower tendency to submit to a tournament (see also Moore and Small (2004)).

7.2 Does Performance Predict who Submits the Piece Rate to a Tournament?

Table 11 shows for women and men, the average piece-rate performance conditional on the task-4 choice.

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 do and do not choose a tournament.⁶⁵ A probit regression of the decision to submit to a tournament conditional on the piece-rate performance reveals that the coefficient on the piece-rate performance is not significantly different from 0.⁶⁶ However, the men who submit to the tournament do have a significantly higher average performance than the men who do not.⁶⁷ A probit regression reveals that the piece-rate performance is significant in predicting the men who submit to a tournament.⁶⁸

Figure 7 shows the propensity of women and men to submit to the tournament for each piece-rate performance quartile. Relative to task 3 we see that the compensation choice appears more sensitive to the participant's performance quartile, and that there is less of a gender gap.

⁶⁵ A t-test comparing the piece-rate performance of women who do and do not select tournament yields $p=0.48$.

⁶⁶ A probit regression of the probability to select tournament as a function of the piece-rate performance reveals a marginal effect of 0.02 (s.e. 0.03, p-value 0.46) evaluated at a piece-rate performance of 11 for women.

⁶⁷ A t-test comparing the piece-rate performance of men who do and do not select tournament yields $p=0.004$.

⁶⁸ A probit regression of the probability of submitting to a tournament on the piece-rate performance reveals a marginal effect of 0.08 (s.e. 0.03, p-value 0.01) evaluated at a piece-rate performance of 11 for men.

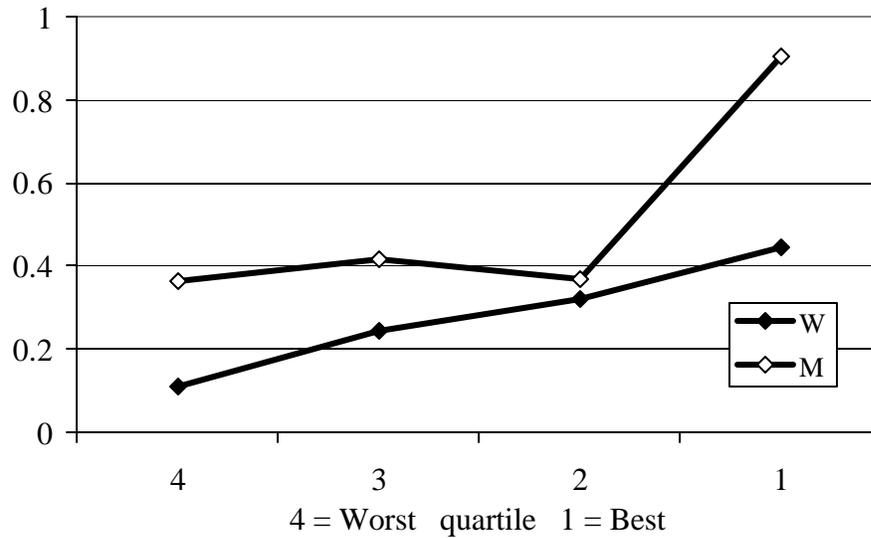


Figure 7: Proportion of Participants who Submit to a Tournament Conditional on Task-1 Performance Quartile.

Of the participants who have higher expected earnings from submitting to the tournament (more than 11 correct answers), significantly more men (14/16) than women (3/12) enter the tournament.⁶⁹ Of the participants who have fewer than 11 correct answers (and less than a 12% chance of winning the tournament) there is no significant difference between the men and the women who submit to the tournament (8/22 and 5/22 respectively).⁷⁰

The observation that conditional on the piece-rate performance men more often submit to a tournament 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). The table presents marginal effects evaluated at a man with 11 correct answers in task-1.⁷¹

The significant gender difference in submitting to a tournament is driven by participants who have more than 11 correct answers. While a probit regression on the sub sample of participants

⁶⁹ p=0.001 using a fisher's exact test.

⁷⁰ p=0.33 with a Fisher's exact test

⁷¹ We evaluate at a piece-rate performance of 11 because this is the point at which the expected payoff is the same from the piece rate as from the tournament.

whose performance exceeds 11 yields a significant female dummy, it is not significant on the subsample of participants with a performance below 11.⁷²

7.3. How do Beliefs on Relative Performance Influence Compensation Choice?

Next we examine if participants' beliefs on piece-rate ranking can help explain the decision to submit to a tournament, however before doing so we study the formation of these beliefs.

7.3.1 Formation of Believed Piece-Rate Rank

If participants are perfectly calibrated and only know the performance distribution and their gender, but not individual performance, then optimal guesses would be first and second rank for men, and second and third rank for women.⁷³ Knowing their actual performance the resulting optimal guesses are reported in Table 13, along with the actual distribution of rank. Neither the actual ranks, nor the optimal guesses differ significantly across gender.⁷⁴

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 can be placed in when controlling for ties.

As seen in Table 14 the elicited distribution of beliefs on relative performance differs both from the actual an optimal distributions. Both, women and men are overconfident and the

⁷² A probit regression of the decision to submit 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 on the female dummy of -0.17 (s.e. 0.14, p=0.23), evaluating the marginal effects at a man who solves 10 problems. For participants who solve 12 or more, the coefficient on the piece-rate performance is 0.03 (s.e. 0.05, p=0.42) and on the female dummy -0.63 (s.e. 0.15, p=0.002), when evaluated at a man who solves 12.

⁷³ 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.

⁷⁴ Chi square tests yield p=0.88 and p=0.54 respectively. Coincidentally the optimal rank conditional on performance is the same for women and men.

elicited rank distribution differs significantly from both the optimal and actual rank distribution.⁷⁵

Table 14: Distribution of Gussed Piece-Rate Rank

	Men		Women	
	Gussed Rank	Incorrect Guess	Gussed 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: Gussed rank out of four.

Once again men have more optimistic beliefs than women about their relative performance, and the distribution of incorrect guesses differs significantly across gender.⁷⁶ To determine whether women and men form different beliefs conditional on their performance, we use an ordered probit to estimate the gussed piece-rate rank as a function of the piece-rate performance and a female dummy, eliminating the few guesses of 4. The results in Table 15 reveal that higher absolute performance causes people to think they have higher relative performance, and that women are significantly less optimistic than men.⁷⁷

Table 15: Ordered Probit of Gussed 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: Guesses of 4 are eliminated leaving 39 women and 38 men.

7.3.2 Do Beliefs Predict who Submits the Piece Rate to the Tournament?

Since the payment from the tournament depends on relative rather than absolute performance, it is natural to examine if, in contrast to absolute performance, beliefs about relative

⁷⁵ Comparing the gussed rank to the optimal gussed 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.

⁷⁶ Chi-square tests yields $p=0.029$ and $p=0.047$, respectively. A t-test comparing the mean beliefs of men and women yields a p-value of 0.04.

⁷⁷ Separate regressions reveal that while for men past piece-rate performance affects the gussed rank that is not the case for women. An ordered probit regression on the gussed piece rate rank (for guesses 1, 2 and 3) as function of the piece-rate performance reveals coefficients of -0.09 (s.e. 0.07, p-value 0.21) for women -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.

performance can explain the gender difference in tournament compensation. Certainly the observation that men are significantly more optimistic about their relative performance suggests that beliefs on ranking may help explain the difference.

Independent of gender more confident participants are more likely to submit to the tournament.⁷⁸ Figure 8 shows, for each guessed piece-rate rank, the proportion of women and men that submit their piece-rate performance to a tournament. Women and men are both about a 60 percentage points more likely to submit to a tournament when they think they are the highest performer in their group, rather than the second highest.

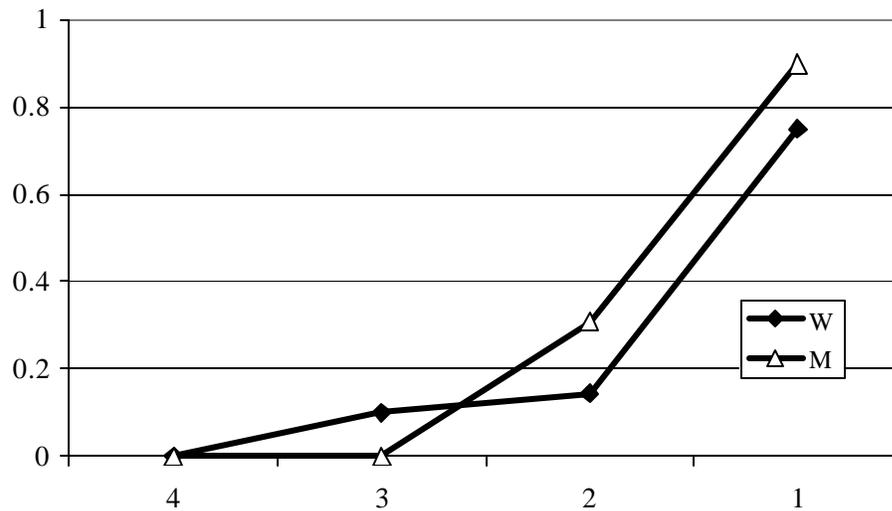


Figure 8: The Proportion of Participants that Submit to the Tournament by Guessed Piece-Rate Rank

In sharp contrast to our task-3 decisions we see that conditional on beliefs the difference between the male and female decision is only about 15 percentage points. A probit analysis confirms that controlling for guessed rank women and men do not differ significantly in their task-4 compensation choice.⁷⁹

⁷⁸ 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.

⁷⁹ Considering only guesses of 1 and 2 (62 participants) the marginal effects are -0.14 on female (s.e.0.11, p=0.13) and -0.31 on guessed rank (s.e. 0.11, p=0.00) evaluated at a man with a guess of 1.

Table 16: Probit of Submitting to the Tournament on Individual Characteristics

	Coefficient	Standard Error	p-value
Female	-0.13	0.11	0.21
Piece Rate	0.00	0.02	0.80
Gussed Piece Rate Rank	-0.32	0.08	0.00

Notes to table: Dependent variable: task-4 compensation choice (1-tournament, 0-piece rate). The table presents marginal effects evaluated at a man with a guess of first and performance of 11. Excluding guesses of 4 the sample is 39 women and 38 men.

While piece-rate performance cannot explain the gender difference in compensation choice this difference can be explained by the participants' relative ranking. Nonetheless it is of interest to determine the relative size of each effect. Eliminating participants with a rank of 4 and controlling for the piece-rate performance women are 33% less likely to submit to the tournament.⁸⁰ Table 16 above shows that including the gussed piece-rate rank the gender effect reduces to 13%. That is 39% of the overall gender effect can be attributed to men and women acting differently upon their beliefs, while 61% can be attributed to them forming different beliefs.

Participants in task 4 decided whether to submit their piece rate result to a tournament or to a piece rate. More men than women submit to a tournament and this is driven by a gender difference among high performance participants. While for men actual piece-rate performance is significantly correlated with the decision to submit to a tournament, it is not for women. However this difference does not explain the gender difference in compensation choice. While overall, both women and men are overconfident about their ability, men are significantly more optimistic than women. When submitting to the tournament this difference in beliefs is sufficient to explain the gender difference in compensation.

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 choice is between a piece-rate versus a tournament payment scheme, and in both cases the decision will depend on the participant's beliefs on relative performance. Furthermore a choice of tournament will in both cases imply that participants receive feedback about their relative performance. The difference between the

⁸⁰ Eliminating guesses of 4 the coefficient on female in table 12 would be -0.327 (s.e. 0.11, p=0.01).

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.

Our study shows that there is a significant gender difference in tournament entry and in beliefs about relative tournament ranking. We want to determine whether gender differences in tournament entry are driven solely by general factors, which are present when participants decide whether to submit the piece rate, or if there are additional gender differences when it comes to performing in a tournament and thinking about one's relative tournament performance.

Our results thus far suggest that the decision to submit to a tournament differs from the decision to enter the tournament and then perform. While for high-performing participants there is a significant gender difference in the rate by which participants submit to the tournament these differences are not significant among low performing participants. In contrast the gender difference in tournament entry is independent of performance. Furthermore, gender differences in belief formation about the piece-rate performance are sufficient to eliminate gender differences in the decision to submit to a tournament, which is not the case when participants decided whether to enter a tournament.

We now estimate the gender difference in tournament entry 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 test whether, controlling for performance, guessed tournament rank, and the decision to submit the piece-rate performance to a competitive payment scheme, there still is a gender difference in the propensity by which participants choose to enter a competitive environment. The participants' task-4 compensation choice serves as a measure for general risk attitude, feedback aversion, etc. The results are reported in Table 17.

Table 17: Probit of Task-3 Compensation Conditional on Performance

	Coefficient	Standard Error	p-value
Female	-0.162	0.100	0.049
Tournament	-0.009	0.012	0.417
Tournament – Piece Rate	0.011	0.015	0.440
Guessed Tournament Rank	-0.117	0.055	0.007
Submitting the Piece Rate	0.258	0.117	0.012

Notes to table: Dependent variable: task-3 compensation choice (1-tournament and 0-piece rate). The table presents marginal effects evaluated at a man solving 13 problems in the tournament and 12 in the piece rate who believes he is ranked first and submits to the tournament.

We see that those who submit to a tournament (task-4) are significantly more likely to enter a tournament (task 3). Including a control for gender differences that are not directly related to tournaments 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 man who submitted the piece rate with a performance of 13 in the tournament and 12 in the piece-rate, were he to be a woman. Not including the task-4 control the overall gender effect was 27.8% (see Table 9), however controlling for the decision to submit the piece rate, the gender effect is reduced to 16.2%, that is about 42% of the remaining gender effect can be explained by general factors, such as risk attitudes, and the residual “competitive” component is 58%.⁸¹

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 16%.

8.2. Gender Differences in Belief Formation

The participants expected outcome of compensation choice are determined by their beliefs about their relative performance. We found that men are more optimistic in assessing their relative tournament 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-

⁸¹ The results are similar when we use task 3 as the performance measure. In this case the coefficient on female drops from -0.279 to -0.160, thus 43% is explained by general factors and the remaining 57% is unexplained.

rate beliefs as an independent measure of how participants form beliefs in general, for example as a measure of overconfidence.

Table 18: Ordered Probit of Believed Tournament Rank on Individual Characteristics

	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
Gussed Piece Rate Rank	0.82	0.28	0.00

Notes to table: Dependent variable: Believed tournament rank. Eliminating guesses of 4 in either the tournament or the piece rate the sample is reduced to 36 women and 36 men.

Table 18 reveals that 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% see Table 9). We want to estimate whether the gender difference is still significant and large once we control for the participant’s 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. A probit of the tournament-entry decision delivers:

Table 19: Probit of Task-3 Compensation Conditional on Believed ranking

	Coefficient	Standard Error	p-value
Female	-0.18	0.11	0.05
Tournament	-0.00	0.01	0.85
Tournament-Piece Rate	-0.01	0.02	0.77
Gussed Tournament Rank	-0.17	0.08	0.00
Gussed Piece Rate Rank	0.14	0.09	0.04
Submitting the Piece Rate	0.47	0.17	0.00

Notes to table: Dependent variable: task-3 compensation choice (1-tournament and 0-piece rate). We eliminate participants with guesses of 4, which leaves us with 37 men and 37 women. The table presents marginal effects evaluated at a man with who guesses first both in the tournament and in the piece rate and who submits the piece rate to the tournament.

The initial gender gap of 28% is reduced by about 36% to 18%, once we control for the decision to submit the piece rate. This suggests that the gender difference in tournament entry is not solely due to a general difference in how participants act upon their beliefs.⁸²

9. Economic Consequences

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

To evaluate expected earnings from the task-3 decision (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 compensation scheme. We consider two extreme ways of assessing the expected costs of tournament entry. In columns 1 and 2 we use the task-2 tournament performance, which is the performance available to participants when they decide whether to enter 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 cost of over entry is higher for men, while the cost of under entry is higher for women. Specifically the low performing men lose by entering the tournament too often, whereas the high performing women lose by not entering enough. Since 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, the total costs are higher for women than for 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) than men, while among low performers total costs are higher for men than for women. Once again the total costs for women are higher than for men.

⁸² 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 first perform under a piece rate and a competitive tournament compensation scheme. Participants then decide in which compensation scheme they want to apply to their next performance. Despite there being no gender differences in initial performance twice as many men as women enter the tournament. Neither performance before nor after the tournament-entry decision can predict the entry decisions of men and women. While men form more optimistic beliefs about their relative ranking (though both genders are overconfident) 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 less “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 examined 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.
- Babcock, Linda and Sara Laschever, *Women Don't Ask: Negotiation and the Gender Divide* by (Princeton University Press, 240 pp., September 2003)
- 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 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.
- Eckel, C., and P. Grossman, "Differences in the Economic Decisions of Men and Women: Experimental Evidence." Forthcoming in *Handbook of Experimental Results*, edited by C. Plott and V. Smith. New York, Elsevier.
- 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), zTree - 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.
- Hewitt, N. M, and E. Seymour, "Factors Contribution to High Attrition Rates Among Science

- and Engineering Undergraduate Majors” *Report to the Alfred P. Sloan Foundation* (1991).
- Hyde, Janet Shibley, Elizabeth Fennema, and Susan J. Lamon, “Gender Differences in Mathematics Performance: A Meta-Analysis,” *Psychological Bulletin* (1990), 107:2, 139-155.
- 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