



How uncertainty and ambiguity
in tournaments affect gender
differences in competitive
behavior

Loukas Balafoutas
Matthias Sutter





How uncertainty and ambiguity in tournaments affect gender differences in competitive behavior

Forthcoming in: European Economic Review (accepted May 2019)

Loukas Balafoutas / Matthias Sutter

May 2019

How uncertainty and ambiguity in tournaments affect gender differences in competitive behavior[#]

Loukas Balafoutas

University of Innsbruck

Matthias Sutter*

Max Planck Institute for Research on Collective Goods Bonn, University of Cologne and

University of Innsbruck

Abstract

Tournament incentives prevail in labor markets. Yet, the number of tournament winners is often unclear to competitors. While it is hard to measure how this uncertainty affects work performance and willingness to compete in the field, it can be studied in a controlled lab experiment. We present a novel experiment where subjects can compete against each other, but the number of winners is either uncertain (but with known probabilities) or ambiguous (with unknown probabilities for different numbers of winners). We compare these two conditions to a control treatment with a known number of winners. We find that ambiguity induces a significant increase in the performance of men who choose to compete, while we observe no change for women. Men also increase their willingness to enter competition in the presence of ambiguity. Overall, both effects contribute to men winning the tournament significantly more often than women under uncertainty and ambiguity. These findings suggest that management should make tournament conditions transparent and information available in order to prevent gender disparities from increasing under uncertainty and ambiguity.

Keywords: Gender, competition, uncertainty, ambiguity, experiment

JEL Codes: C91; D03; D09

Forthcoming in: European Economic Review (accepted May 2019)

[#] We received very helpful comments from Stefan Trautmann and three referees. We also wish to thank Brent J. Davis, who has made valuable contributions to this work and was originally part of the team of authors. Financial support from the University of Cologne (through the Excellence Initiative and the Hans-Kelsen-Prize) is gratefully acknowledged.

* Corresponding author. Max Planck Institute for Research on Collective Goods. Kurt-Schumacher-Strasse 10, D-53113 Bonn. matthias.sutter@coll.mpg.de

1 Introduction

Tournament based incentive schemes are ubiquitous in the labor market as a means to induce high effort of workers. Moreover, many companies across practically all industries use them for determining promotions or wages. Of course, the aim is to assign the best people to the right jobs and to reward outstanding performance with high wages. The fulfillment of this aim should not be obstructed by the gender of competitors. Yet, there are still large gender differences with respect to wages and promotion opportunities, especially with respect to the fraction of women in high level executive positions and at high income levels (Weichselbaumer and Winter-Ebmer, 2007; Blau and Kahn, 2017). For instance, at the beginning of 2017, women accounted for less than 6% of CEOs within the S&P 500 and the Fortune 500.¹ Although some of these differences can be explained by sorting effects, there is still a persistent gender gap in firm-specific pay (Card et al., 2016) and the gender pay gap remains even after parenting and family concerns are taken into account (Angelov et al., 2016). So, what are the sources of these gender differences and what could management do to achieve a better gender balance in tournaments?

Besides long-standing demand side explanations such as, e.g., discrimination against women, there is mounting evidence that the supply side of labor markets also contributes to the gender gap in wages and promotions. There has been an ever growing literature showing that women often perform worse than men under competitive pressure (pioneered by Gneezy et al., 2003; see also Gneezy and Rustichini 2004; Antonovics et al., 2009; Gill and Prowse, 2014; Almås et al., 2016) and that women shy away from competition much more often than men do (seminal paper by Niederle and Vesterlund, 2007; see also Booth and Nolen, 2012; Datta Gupta et al., 2013; Buser et al., 2014; Brandts et al., 2015; Sutter and Glätzle-Rützler, 2015; Flory et al., 2015; Saccardo et al., 2018).² Since firms often use tournament payment and promotion schemes (Eriksson, 1999; Bognanno, 2001), a lower inclination of women towards competition can help explain to some extent why they are less often promoted and end up with lower wages.³ Providing empirical evidence in this vein, McGee et al. (2015) show on the basis of US data

¹ Data for the S&P500 CEOs is from, *Women CEOs of the S&P500*, available at www.catalyst.org (accessed May 19, 2017); Data for the Fortune 500 list is from *Female Fortune 500 CEOs Are Poised to Break This Record in 2017*, available at <http://fortune.com/2016/12/22/female-fortune-500-ceos-2017/> (accessed August 19, 2017).

² Recent research shows that affirmative action policies can be successful in increasing women's willingness to compete, with only limited harmful side-effects (Balafoutas and Sutter, 2012; Villeval, 2012; Niederle et al., 2013; Calsamiglia et al., 2013; Sutter et al., 2016; Kölle, 2017; Maggian and Montinari, 2017; Banerjee et al., 2018). Additional recent research on gender and competitiveness studies competitive behavior in adolescence (Dreber et al., 2011; 2014; Sutter and Glätzle-Rützler, 2015) or the impact of family background (Almås et al., 2016), hormones (Wozniak et al., 2014), advice (Brandts et al., 2015), stress (Buser et al., 2017), prize structure (Petrie and Segal, 2015), information and feedback (Ertac and Szentos, 2011; Wozniak et al., 2014), and team decision-making on competitive behavior (Healy and Pate, 2011, Dargnies, 2012).

³ A somewhat related factor contributing to the gender gap is that there appear to be differences in salary negotiations between men and women (Leibbrandt and List, 2014; Card et al., 2016; Rigdon, 2016).

that women are less likely than men to be remunerated with a competitive compensation package and that this can explain part of the wage gap between the two genders.

Interestingly, and only with a few exceptions discussed below, the extensive (lab and field) experimental literature on gender differences in competitive behavior has mainly investigated situations where subjects who may compete in a tournament are fully informed of the number of potential competitors and the number of winning positions in the tournament. Of course, competition on labor markets rarely meets such a stylized situation. In reality, it is highly unlikely that individuals in the labor market know the environment surrounding a (promotional) tournament with perfect information. For example, consider an associate at a law firm who knows that he or she competes in a tournament with other lawyers at the firm to be promoted. Often this individual does not know how many individuals will be promoted by the firm, and therefore is unaware of the *ex ante* likelihood of being promoted even if the number of competitors were known. How does the uncertainty in such a situation affect the associate's incentives to perform and does it influence the associate's likelihood to seek promotion in such a vague environment? These questions are at the core of our paper. Answering them will be useful for the practice of management, as it will clarify which conditions are more or less conducive to a gender-balanced composition of winners in tournaments.

We present an experiment that – while building on the classical design of Niederle and Vesterlund (2007) to measure the willingness to compete – is novel in the sense that it introduces uncertainty and ambiguity into a tournament with respect to the number of winners and therefore allows examining how uncertainty and ambiguity affect (i) workers' willingness to enter a competition and (ii) their performance in that competition. With our design, we can measure whether the gender differences with respect to performance under competition and willingness to compete depend on whether the tournament environment features uncertainty and ambiguity. Note that our approach of implementing uncertainty or ambiguity about the number of winners in case of a known number of potential competitors is the mirror image of having uncertainty or ambiguity about the number of competitors in case of a known number of winners. For our purposes, the original design of Niederle and Vesterlund (2007) was an ideal starting point that made it very easy to implement our novel treatment variations.

We are not aware of any study that can measure the effects of uncertainty or ambiguity on both the willingness to compete and the performance of potential competitors. Yet, there are a few related studies that implemented tournaments with imperfect information and examined how the latter affected subjects' willingness to compete. Most notably, Flory et al (2015) examine in their field experiment the application rates for job advertisements where they alter

the payment scheme (making it more or less competitive), showing that fewer women apply if a portion of the pay is based on a competitive payment scheme (relative performance) or is uncertain (when it depends on whether the work led to a published paper for the authors), compared to a flat hourly wage payment scheme. Based on these findings, they raise the point that “*high wage uncertainty may be just as important as competition per se in affecting job-entry choices and the gender gap*” (p. 125). In their experiment there is, of course, ambiguity about the number of applicants, but they don’t have a control treatment where this ambiguity is resolved – making it impossible to investigate the pure effects of ambiguity on the gender difference in the willingness to compete. Moreover, Flory et al. (2015) were interested in application rates, but not in competitors’ performance, an issue that we can also study in our design.⁴

In a field experiment using LinkedIn, Gee (2018) varies the information available to potential job applicants. When they have information regarding how many individuals have already clicked on the application link, there is some evidence that this leads to an increase in the likelihood of women (but not men) completing the job application. Such behavior would be consistent with women being more averse than men towards an ambiguous tournament environment. Moreover, providing this information leads to an increase in the likelihood of women applying for a male-dominated job. Hence, it seems that resolving ambiguity has different effects on men and women with respect to their willingness to compete. Again, Gee (2018) differs from our work since performance is not a relevant measure in her experiment. Less related to our paper is Leibbrant and List (2014) who find that women negotiate their wage more often when the uncertainty on whether wages are negotiable or not is removed, indicating again an influence of ambiguity on the labor market behavior of men and women. Generally speaking, these papers show that the level of information available in labor markets affects employees’ behavior in competitive environments. With respect to the willingness to compete, Wozniak et al. (2014) and Ertac and Szentes (2011) have shown that repeated exposure to competition – and the information drawn from it – can reduce the gender gap in the willingness to compete when the number of winners is known for sure. Hence, information matters.

Different from this line of the literature, our paper examines gender differences in the willingness to compete and in work performance under different tournament environments when the information about the number of winners in a one-shot competition may be

⁴ On the issue of performance, there is abundant experimental literature examining the relationship between prize structure and performance in tournaments (Orrison et al., 2004; Harbring and Irlenbusch, 2008; Freeman and Gelber, 2010; Müller and Schotter, 2010; Dutcher et al., 2015). However, in all of these studies, the number of prizes is always known with certainty.

incomplete. More precisely, while we keep the number of potential competitors constant and known, we vary the information about the number of winners as follows: (1) We implement a control treatment where the number of tournament winners is certain and known. (2) We design a tournament where the number of winners is uncertain ex-ante, but the probabilities of how many winners there will be are known. (3) We have an ambiguous treatment where the number of winners is uncertain and the probabilities of how many winners there will be are unknown. Thus, we compare a tournament with certainty regarding the number of winners against a tournament with uncertainty and against a tournament with ambiguity. We have argued that uncertain or ambiguous environments are important to study and that they can often better capture the characteristics of labor markets. Our study may also have policy implications if we find that uncertainty and ambiguity affect men and women differently on labor markets, which would suggest that companies can influence male and female behavior (such as the willingness to compete or performance incentives) by changing the available information to competitors. The latter might be a non-expensive and much less invasive alternative to institutionalized interventions such as affirmation action programs.

We find that uncertainty and ambiguity cause both men and women to increase their likelihood to enter competition, but that the effect is only significant for the ambiguous tournament in the case of men. The gender gap in the willingness to compete increases on average (but not significantly) with uncertainty or ambiguity, compared to the control treatment with certainty. Moreover, men increase their performance significantly – conditional on having entered the competition – when the number of tournament winners becomes uncertain or ambiguous, while the performance of women decreases on average, albeit not significantly. In combination, the effects of uncertainty and ambiguity on performance *and* willingness to compete lead to our final finding that men win the tournament significantly more often than women whenever either uncertainty or ambiguity is involved. Hence, given that competition on labor markets is very often characterized by uncertainty and ambiguity, the previous experimental studies with perfect information and certainty about the number of competitors and winners seem to have measured only a lower bound for the gender differences in competitive behavior and the likelihood with which women win the tournaments. This insight has implications for managerial practice as it suggests that more precise information about the conditions in a tournament will promote a more balanced gender composition of winners. Avoiding uncertainty and ambiguity limits the gender gap in competitive performance and preferences.

Our paper proceeds as follows. Section 2 introduces the experimental design and our expectations about male and female behavior in the different treatments. Section 3 presents the experimental results, and section 4 concludes the paper by discussing our main results and putting them into perspective.

2 The Experiment

2.1 Experimental design and procedures

We build our design in the control treatment on Niederle and Vesterlund (2007), and then introduce uncertainty and ambiguity in two additional treatments. All three experimental treatments share the following characteristics. At the beginning of the experiment, groups of four (composed of two men and two women) were formed randomly and remained unchanged throughout the experiment.⁵ Subjects knew that there would be four stages in the experiment, but only received instructions for the next stage after completing the previous one (instructions are provided in the online appendix). In Stages 1-3, subjects worked on adding as many sets of five two-digit numbers as possible within four minutes. Subjects were not allowed to use calculators, but scratch paper and pens were provided to them. After each attempted solution to an addition problem, subjects were informed if their answer was correct or not, and they received a new problem. The four stages looked as follows:

Stage 1 – Piece Rate: Each subject received €1 for each problem correctly solved within the four minutes' time frame. Subjects were informed how many addition problems they had solved correctly at the end of this stage. Prior to Stage 1, subjects had one minute as a practice round to familiarize themselves with the task.

Stage 2 – Tournament: In this stage, the four subjects within a group competed against each other. The group member who solved the most addition problems in four minutes received €4 per correct problem solved, while the other three group members received no payment. Ties were broken randomly. Subjects were informed how many addition problems they had solved correctly at the end of this stage, yet they were only informed at the end of the experiment

⁵ Niederle and Vesterlund (2007) do not explicitly mention the gender composition in a group, but inform subjects that all group members are seated in the same row, which implies that the gender composition (of two men and two women per group) could be inferred very easily. The difference in our design is that we explicitly informed subjects of the gender composition within a group in the instructions, but without showing who the group members were. This method, also used in Balafoutas and Sutter (2012), has the disadvantage that it could potentially prime subjects on gender or make stereotypes salient, but it has the advantage that it does not reveal who the group members are, thus better preserving anonymity. In any case, it should be noted that the two methods are equivalent in terms of the information that subjects have at their disposal. They are aware of the gender composition within a group, such that any treatment differences in behavior cannot be due to different beliefs about gender composition.

whether they had won or not. Group size and gender composition of the group were revealed to subjects prior to the start of Stage 2.

Stage 3 – Tournament Choice: At the beginning of this stage each group member chose whether he or she wanted to solve the addition problems under a piece-rate scheme as in Stage 1 or under a tournament scheme. If the subject chose the tournament scheme, then that subject's Stage 3 performance was compared to his or her group members' Stage 2 performances.⁶ Any ties were broken randomly. The rules for determining tournament winners were varied across our three treatments as follows:

- 1. Certain Number of Winners (CERTAIN):** There were two winners in Stage 3 for certain. The two⁷ group members with the largest number of correct additions won the tournament. Subjects who won the tournament in Stage 3 received €2 per correct problem, while subjects who lost the tournament received no payment for this stage.
- 2. Uncertain Number of Winners (UNCERTAIN):** There was an uncertain number of winners. Specifically, with 1/3 probability there would be either one, two, or three tournament winners. If there was one winner, the person with the largest number of correctly solved addition problems received €4 for each correct solution. If there were two winners, the same rules as in CERTAIN applied (€2 per correct solution for each of the two winners). If there were three winners, a subject had to solve more problems than at least one of his or her group members in order to win, and each of the three winners received €1.33 for each correct problem solved. All of these design details were known to subjects. Note that in UNCERTAIN the expected number of winners and payment per correct addition problem was the same as in CERTAIN, but uncertainty about this number was added to the environment. Subjects were only informed about the number of winners and whether they won the tournament at the end of the experiment.
- 3. Ambiguous Number of Winners (AMBIGUOUS):** This treatment was similar to the UNCERTAIN treatment, except that subjects did not know the likelihood of there being one, two, or three tournament winners – in contrast to UNCERTAIN. Subjects still knew

⁶ The attractive feature of this design choice of Niederle and Vesterlund (2007) is that subjects are competing against other subjects' (past) performance under a tournament scheme, while at the same time each subject's entry choice does not affect other participants as the pool of competitors is fixed. Hence, a subject's entry decision has no externalities on others and the entry decision cannot depend on the subject's expectation about the potential competitors' entry decisions.

⁷ Recall that in Stage 2 we had only one winner per group. Changing the tournament from having one winner (in Stage 2) to two winners (in Stage 3) in CERTAIN means that even participants in CERTAIN face a change in the rules between the two stages – for which reason Stage 3 is not a simple replication of Stage 2 for them. Since subjects in UNCERTAIN and AMBIGUOUS face a change from Stage 2 to Stage 3 for sure (see description of these treatments), we think this design avoids a potential confound which would arise if CERTAIN was a mere replication of Stage 2, while the other two treatments were not.

how the payments were conditional on the number of winners. Hence, the only change compared to the UNCERTAIN treatment was that the probabilities were unknown, meaning that ambiguity was additionally introduced into the tournament. Actually, we used the same probabilities as in CERTAIN, but subjects were unaware of this. Again, subjects were only informed about the number of winners and whether they won the tournament at the end of the experiment.

We used a between-subjects design, meaning that each subject participated in only one of the three treatments. After Stage 1 we elicited beliefs on which quartile subjects believed they ranked within the entire session (of 20 subjects each). After Stage 2 we elicited beliefs on where subjects thought they ranked within their group of four. Both belief elicitation questions were unannounced to subjects before they occurred. Subjects received €1 for each correct belief. Feedback about beliefs was provided only at the end of the experiment.

In Stage 4, risk and ambiguity attitudes were elicited via the Ellsberg two-color choice task (Ellsberg, 1961). Subjects were presented with 20 choices between a sure amount of money and drawing a ball from a virtual bag (with balls of two different colors) for a chance to receive €6. The sure amount increased from €0.30 to €6 in increments of €0.30 (see also appendix). When eliciting risk attitudes, subjects were told that when drawing from the virtual bag there was an equal chance to win (by drawing an orange ball) or lose (by drawing a white ball). A subject is classified as risk averse if he or she prefers an amount below €3 over the lottery, as risk loving if she prefers the lottery over an amount larger than €3, and as risk neutral otherwise. In the ambiguity experiment, subjects were not told the distribution of white and orange balls, but were informed that each distribution was equally likely (which keeps the expected winning probability at 50%, like in the risk elicitation task).⁸ Again, subjects won €6 if they drew an orange ball. A subject's attitude towards ambiguity can be measured as the difference between the certainty equivalent in the risk elicitation experiment (CE_r) and the certainty equivalent in the ambiguity aversion experiment (CE_a). A subject is classified as ambiguity averse if $CE_a < CE_r$, ambiguity loving if $CE_a > CE_r$, and ambiguity neutral otherwise. We control for risk and ambiguity aversion since our experimental treatments vary the degree of risk and ambiguity to which the experimental subjects were exposed.

⁸ For practical reasons ambiguity was implemented by 20 different and equally likely distributions of balls. This might have reduced overall ambiguity aversion if subjects interpret the prospect as a compound lottery (Trautmann and van de Kuilen, 2015, p 94). However, for our purposes it is important that ambiguity attitudes elicited with compound lotteries are typically closely related to those elicited by fully unknown distributions, especially as we use an Ellsberg-type format in our design (e.g., Halevy, 2007). We therefore believe that ambiguity in the elicitation task and in the experimental treatment AMBIGUITY are perceived as similar by the subjects. Actually, also in the game, subjects may have implicitly assumed that all distributions are equally likely.

Subjects were paid for one of the stages 1, 2, or 3 (chosen randomly), their beliefs, one randomly selected decision from the 20 decisions in the risk task, and one randomly selected decision from the 20 decisions from the ambiguity task in Stage 4. The order in which the risk and ambiguity tasks were presented in Stage 4 varied between sessions (yet we found no order effects). Subjects were informed about the actual payments only at the end of the experiment.

The experiment was run with the z-tree software (Fischbacher, 2007) in the EconLab at the University of Innsbruck. A total of 240 subjects participated in the experiment and they were recruited via h-root (Bock et al., 2014). CERTAIN was run with 48 subjects (24 men and 24 women), while UNCERTAIN and AMBIGUOUS had 96 subjects each (48 men and 48 women). Sessions lasted approximately 60 minutes and the average payment per subject was €19.02.

Before proceeding to the hypotheses, we would like to note that a recent paper by Gillen et al. (2018) has highlighted the perils of measurement error in economic experiments. Taking the design of Niederle and Vesterlund (2007) as an example, Gillen et al. (2018) show that if one uses multiple measures of confidence and risk aversion to reduce this error, the residual gender gap in the willingness to compete disappears. Given that our experiment was not designed to tackle this issue, we do not have multiple measures on any of our key explanatory variables like confidence or attitudes towards risk and ambiguity. However, this important issue must be kept in mind in future research on this topic.

2.2. Hypotheses

We expected to replicate the stylized finding of men choosing competition more often than women in CERTAIN. Many papers have found a ratio in the willingness to compete of men and women close to 2:1 (e.g., Niederle and Vesterlund, 2007; Sutter and Glätzle-Rützler, 2015), and we expected a similar ratio. In many of the experiments that use adding up numbers as experimental tasks, men perform slightly better than women, but frequently the difference is insignificant. So, we expected performance levels of men and women to be fairly close to each other.

Given a large literature showing that women are typically more risk averse than men (e.g. Croson and Gneezy, 2009; Borghans et al., 2009; Charness and Gneezy, 2012), and in addition assuming that a higher degree of risk aversion leads to a lower willingness to enter competitive environments (Niederle and Vesterlund, 2007), we expected that the gender gap in the willingness to compete should be larger in UNCERTAIN than in CERTAIN. It is a priori less clear what to expect with respect to performance levels of men and women in

UNCERTAIN. While we expected fewer women to enter competition, *conditional on having chosen competition (or opted out)* we did not expect performance differences in comparison to CERTAIN. In combination, this would imply an overall lower performance level of women in UNCERTAIN than in CERTAIN, under the assumption that those women who opt into competition perform better than those who opt out (which is an empirical regularity that is typically found in the literature).

Concerning gender differences in ambiguity aversion, the evidence itself is ambiguous. Some studies find that women are more ambiguity averse (e.g., Powell and Ansic, 1997), a majority of studies find little or no difference (e.g., Borghans et al., 2009; Sutter et al., 2013), and a few studies report less ambiguity aversion of women in some domains (e.g., Schubert et al., 2000). Assuming that there is no gender difference in ambiguity aversion in our pool, our expectations for AMBIGUOUS would match those for UNCERTAIN. If we found women to be more ambiguity averse, the gender gap in the willingness to compete should increase from UNCERTAIN to AMBIGUOUS, increasing also the gap in average performance between men and women.

Another important factor that might affect tournament entry rates across treatments and gender relates to participants' beliefs about their rank in the tournament. It is straightforward to see that subjects who believe to be the best performer in their groups should enter the tournament in any case and in all treatments, because winning the tournament yields in every treatment a larger amount per correct solution than staying out and taking the piece rate of €1 per correct solution. Subjects who believe they will rank second in their group should be less likely to compete in UNCERTAIN and AMBIGUOUS than in CERTAIN, because in UNCERTAIN and AMBIGUOUS there is a chance that ranking second is not enough to ensure winning the tournament. On the contrary, a subject who believes she will rank third in the group should be more likely to enter the tournament in UNCERTAIN and AMBIGUOUS than in CERTAIN, because in the former two treatments there is a positive probability of a third-placed participant winning the tournament. Finally, subjects who expect to rank last within their group should choose the piece rate irrespective of treatment.

Whether subjects are more or less likely to compete in AMBIGUOUS than in UNCERTAIN is less clear, since it depends on their subjective beliefs about the probability of two or three winners in the tournament. One way around this is to impose some structure on the form of ambiguity preferences. It is straightforward to show that, using the alpha-maxmin model (Ghirardato et al., 2004), participants with an expected rank of 2 should be more likely

to enter the tournament in UNCERTAIN than in AMBIGUOUS, while the opposite is the case when the expected rank is 3.⁹

However, hypotheses for tournament entry decisions by expected rank based on the above considerations cannot be tested with our dataset, because we have neither elicited subjects' subjective probabilities in the ambiguity treatment, nor their expected rank in the tournament of Stage 3 (since not all subjects participated in the tournament in that stage). What we are able to do is to account in the regression analyses for the role of participants' beliefs about their tournament performance, using the expected rank in Stage 2 as a proxy for beliefs about Stage 3. As we will show in section 3.1, these beliefs are independent predictors of competition entry decisions, but they do not interact significantly with the uncertainty or the ambiguity treatment. Hence, the effect of expected rank does not appear to differ across treatments, making it an unlikely driver of potential treatment differences in tournament entry.

3 Results

3.1 Willingness to compete

Figure 1 shows the relative frequency of entry into competition, broken down by gender and treatment. Consistent with the literature and with our expectations, men enter competition at higher rates than women in all treatments (CERTAIN, $p=0.08$; UNCERTAIN, $p<0.01$; AMBIGUOUS, $p<0.01$; χ^2 tests).¹⁰ In CERTAIN, the entry ratio is exactly 2:1, which matches previous findings (Niederle and Vesterlund, 2007; Balafoutas and Sutter, 2012).

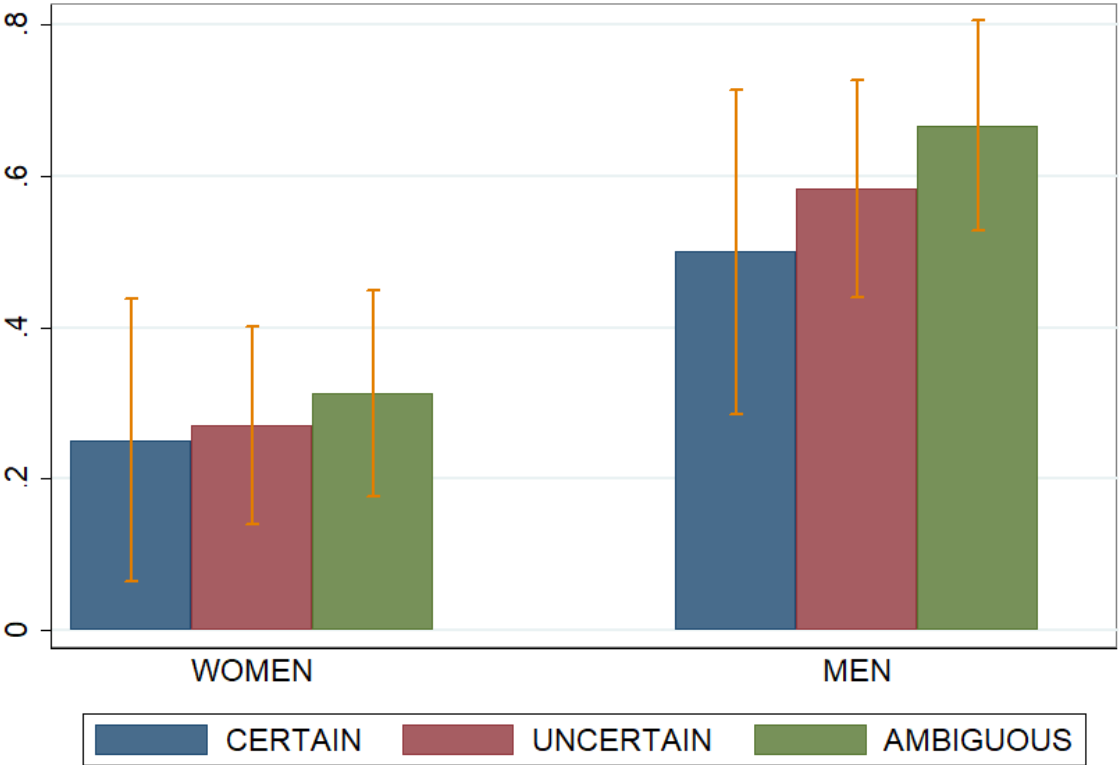
Contrary to our expectations, women slightly increase their likelihood to compete in UNCERTAIN and AMBIGUOUS, moving up the entry-rate from 25% in CERTAIN to 27% in UNCERTAIN and 31% in AMBIGUOUS. None of these differences are significant, however. Men also increase their likelihood to compete from 50% in CERTAIN to 58% in UNCERTAIN and 67% in AMBIGUOUS. The difference between CERTAIN and AMBIGUOUS is insignificant using a two-sided χ^2 test ($p=0.17$), although we will show in the regression analysis that controlling for other variables leads to a significant effect of ambiguity on men's willingness to compete. As a result of the above patterns, the gender gap in the

⁹ These calculations follow from comparing the expected prize in the Stage 3 tournament across treatments, conditional on expected rank in Stage 3. The value of alpha used for the calculations is 0.56, following Dimmock et al. (2005), but the results remain qualitatively the same if we follow Ahn et al. (2014) who estimate an alpha of 0.52. We assume that the performance rank itself is not ambiguous, and that beliefs about performance ranks are degenerate. We also note that using the maxmin model of Gilboa and Schmeidler (1989) leads to the prediction that entry rates should be lower in UNCERTAIN than in AMBIGUOUS for expected ranks of 2 and 3.

¹⁰ We note that all p-values reported in the analysis refer to two-sided tests.

willingness to compete increases from 25 percentage points in CERTAIN to 31.2 percentage points in UNCERTAIN and 35.4 percentage points in AMBIGUOUS.

Figure 1. Relative Frequency of Tournament Entry by Gender and Treatment (including 95% confidence intervals)



A multivariate analysis of entry into competition is provided in Table 1, which shows marginal effects of Probit regressions. The dependent variable is choosing to enter the tournament in Stage 3. The right hand side variables are the following: *Male* is a dummy variable equal to 1 for men and 0 for women; *Uncertain* is a dummy variable equal to 1 when the treatment is UNCERTAIN and 0 otherwise; *Ambiguity* is a dummy variable equal to 1 in treatment AMBIGUOUS and 0 otherwise; *Belief2* is a subject’s reported belief about his or her rank in the Stage 2 tournament (recall that lower ranks indicate better performance); *Correct2* is a measure of performance in a competitive situation, namely the number of problems the subject solved correctly in Stage 2; *Risk Measure* and *Ambiguity Measure* are the subjects’ elicited risk and ambiguity measures in Stage 4. The risk measure ranges from 0 (most risk seeking) to 1 (most risk averse), with 0.5 indicating risk neutrality. The ambiguity measure has positive values for ambiguity aversion, zero for ambiguity neutrality and negative values for ambiguity loving. Finally, *Economics* is a dummy variable equal to 1 for those subjects who were studying Business or Economics at the time of the experiment (approximately 35% of the

sample). Specifications 2 and 3 also include interaction terms of *Uncertain* and *Ambiguity* with gender, with both interaction terms having no significant effect on tournament entry.

The regression estimates in Table 1 confirm our graphical impression regarding tournament entry. Men enter competition significantly more often than women: this is captured by the significant positive coefficient on *Male* in models (1) and (2), and by the significant joint coefficient of *Male* plus all relevant interaction terms where *Male* is included in model (3) ($p < 0.01$). In terms of treatment effects, introducing uncertainty is not enough to lead to a significant increase in competition entry rates compared to the baseline as shown by the lack of significance on the *Uncertain* coefficient. However, treatment AMBIGUOUS does have a significant impact on competition entry as seen by the significant coefficient for *Ambiguous* in model (2). Specifications (4) and (5) present the regressions for each gender separately, and reveal that the effect of ambiguity is driven by men: moving from the certain to the ambiguous environment significantly increases tournament entry rates among men, but not among women. All of these effects lead to our first result.

Result 1: Introducing uncertainty and ambiguity increases competition entry rates by men and by women, contrary to expectations, but the difference in entry rates across treatments is not significant with the exception of tournament entry by men in the AMBIGUOUS treatment. The gender gap in entry rates becomes larger with uncertainty and ambiguity in absolute terms, but the increase is not significant.

In terms of additional control variables, we find that subjects enter competition more often when they solve more problems in Stage 2 (*Correct2*), when they believe they ranked better within their group in Stage 2 (i.e., when their perceived rank *Belief2* is lower), and when they are enrolled in an economics study program.¹¹ Results are robust when we replace *Belief2* with dummies for the different possible guessed ranks and when we use Stage 1 instead of Stage 2 performance as our measure of ability. Moreover, in a regression not shown here we have estimated a version of specification (3) in which we add interaction terms of gender with beliefs and performance. These interaction terms are insignificant and including them does not lead to any substantial change in our findings. Finally, we have estimated versions of specifications (2) and (3) in which we interact *Belief2* with the two treatment dummies. These interaction terms

¹¹ Consistent with previous studies (see Croson and Gneezy, 2009, for a short summary) we find that men are much more confident about their performance than women. In Table A1 in the Appendix we report elicited beliefs about one's rank within a group in Stages 1 and 2. Men report significantly lower (i.e., better) expected ranks than women for their performance in Stage 1 and in Stage 2 ($p < 0.01$ for each stage, χ^2 tests).

are insignificant and their inclusion does not affect any of the other results. Given our discussion in 2.2 regarding the possible role of beliefs in shaping treatment differences, it is quite reassuring that beliefs about performance in Stage 2 do not account for the observed treatment effects.

Table 1. Regression Estimates for Entry into Competition

	(1)	(2)	(3)	(4)	(5)
	Pooled	Pooled	Pooled	Men	Women
<i>Male</i>	0.319*** (0.061)	0.251*** (0.069)	-0.405 (0.285)		
<i>Uncertain</i>	0.057 (0.093)	0.127 (0.105)	0.044 (0.134)	0.167 (0.140)	0.069 (0.116)
<i>Ambiguous</i>	0.124 (0.092)	0.214** (0.103)	0.130 (0.136)	0.281** (0.136)	0.120 (0.117)
<i>Belief2</i>		-0.264*** (0.052)	-0.293** (0.053)	-0.304*** (0.071)	-0.182*** (0.062)
<i>Correct2</i>		0.03** (0.012)	0.032** (0.013)	0.019 (0.015)	0.031** (0.014)
<i>Uncertain x Male</i>			0.077 (0.215)		
<i>Ambiguous x Male</i>			0.170 (0.211)		
<i>Risk Measure</i>			-1.248*** (0.373)		
<i>Risk Measure x Male</i>			1.077** (0.522)		
<i>Ambiguity Measure</i>			-0.698* (0.365)		
<i>Ambiguity Measure x Male</i>			1.119** (0.537)		
<i>Economics</i>			0.137* (0.081)		
Observations	240	240	231	120	120

Robust standard errors in parentheses. ***, **, * indicates significance at the 1%, 5%, and 10% level, respectively. Models are Probit with marginal effects reported. Dependent variable is choosing to enter competition in Stage 3. Nine observations are dropped in Model 3 as nine subjects had inconsistent choices in either the risk or ambiguity measurement in Stage 4.

Model (3) includes the measures of risk and ambiguity aversion elicited in Stage 4 of the experiment, as well as interaction terms of these measures with gender. The coefficient on *Risk Measure* is significant and in the expected direction, i.e., a higher degree of risk aversion leads to a lower likelihood of competition entry. This effect is restricted to female subjects, as seen by the significant interaction term of the risk measure with *Male* that leads to an insignificant effect for men (*Risk Measure + Risk Measure x Male*, $p=0.64$, χ^2 test). The ambiguity measure has a weakly significantly negative coefficient, suggesting that more ambiguity averse female subjects are also less likely to choose the tournament payment, while the effect is reversed and becomes insignificantly positive in the case of men (*Ambiguity Measure + Ambiguity Measure x Male*, $p=0.28$). Hence, it appears that gender differences in the willingness to enter competition are largely driven by gender differences in risk preferences and in particular by the fact that risk and ambiguity aversion affect choices among female, but not among male participants.¹² In Table A3 in the Appendix we offer more support for this pattern by presenting separate regressions by treatment. These regressions feature an insignificant coefficient for *Male*, coupled with a significant impact of risk aversion on the choices of women (significant coefficients for *Risk Measure* and *Ambiguity Measure*), but not on the choices of men given the insignificant joint coefficients of the two measures and their interactions with *Male*.

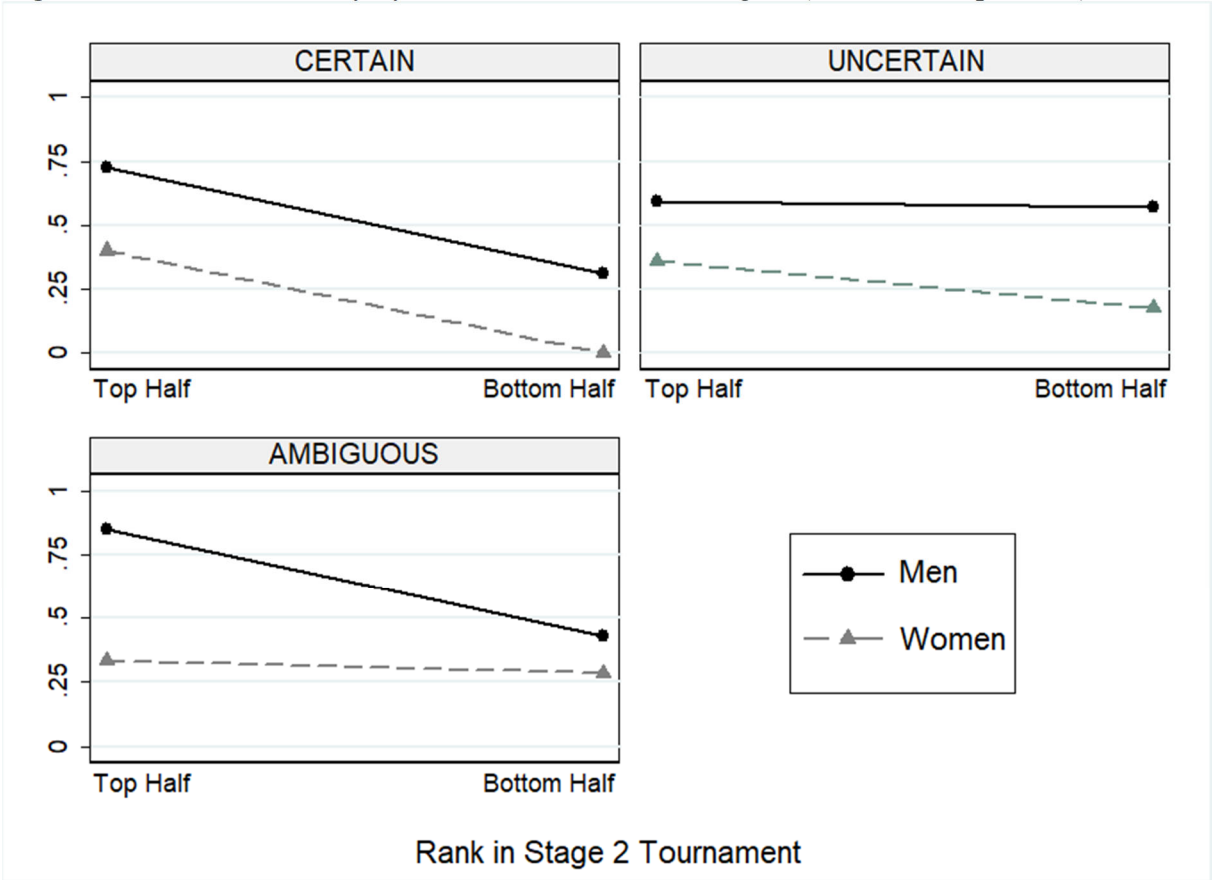
Introducing uncertainty and ambiguity into the tournament environment may not only affect the absolute frequency of tournament entry, but also the entry rates conditional on individual performance. Figure 2 shows tournament entry conditional on a subject's performance rank in Stage 2. As we have two winners in expectation, Figure 2 is grouped by subjects who either rank in the top half (rank 1 and 2) or the bottom half (ranks 3 and 4).

It is straightforward to see that men enter competition more often than women at all performance ranks and in all three treatments. Looking at the top-half performers, the difference in entry rates between men and women is impressive and large in all treatments. In CERTAIN, it is 33 percentage points (72.7% for men in the top half, compared to 40.0% for women in the top half; $p=0.10$, χ^2 test). In UNCERTAIN, it is 23 percentage points (59.3% vs. 36.0%; $p=0.09$), and in AMBIGUOUS it is most pronounced with 52 percentage points (85.2% vs. 33.3%; $p<0.01$). Among subjects who perform in the bottom half, men enter the competition more than women in CERTAIN (30.8% vs 0%, $p=0.07$), UNCERTAIN (57.1% vs. 17.4%, $p=0.01$), and AMBIGUOUS (42.9% vs. 28.6%, $p=0.33$).

¹² For completeness, Table A2 in the Appendix summarizes the risk and ambiguity measures for men and women. We find that women are more risk averse than men, while there are no gender differences in ambiguity attitudes and no treatment differences in the risk or ambiguity measure.

It follows from the above analysis that having more information regarding the tournament environment (i.e., being in CERTAIN as compared to the other two treatments) changes the pool of people who enter the tournament, and as such better identifies the most talented individuals in groups. Figure 2 shows that across treatments there are hardly differences in the entry rates of the top performers. This holds for both men and women, although women always enter much less frequently. However, low performing men and women are more likely to enter the tournament in the uncertainty and ambiguity treatments: the rate of low performers who compete in Stage 3 (pooled for both genders) approximately doubles, from 18.2% in CERTAIN to 36.4% in UNCERTAIN and 35.7% AMBIGUOUS. This can be quite costly for low performing participants, who would be better off selecting the piece rate rather than entering the tournament – and losing. Moreover, while there are no associated costs with selecting winners in the lab, this increase of low-performers entering competition may impose significant costs outside of the lab on firms screening possible candidates for jobs or promotion.

Figure 2. Tournament Entry by Performance Rank in Stage 2 (Relative Frequencies)



3.2 Performance

Table 2 reports mean performance (correctly solved exercises) of men and women, by treatment and stage, as well as Mann-Whitney tests on performance differences compared to the control treatment. Recall that Stages 1 and 2 did not differ across treatments. The data in Panel (a) of Table 2 show that performance in Stages 1 and 2 does not vary significantly across treatments. Moreover, there is no significant gender difference in Stage 1 performance, neither when we pool across all treatments (8.13 vs. 7.80, $p=0.69$, two-sided Mann Whitney test) nor when we look at individual treatments. The same applies to Stage 2 where there is also no gender difference in performance (8.97 for men vs. 8.63 for women, $p=0.38$).¹³ However, both men and women increase their performance from Stage 1 to Stage 2 significantly (men: 8.97 vs. 8.13, $p<0.01$; women: 8.63 vs. 7.80, $p<0.01$, Wilcoxon signed-rank tests).

When we compare performances in Stage 2 and Stage 3, we see no significant differences in CERTAIN and in UNCERTAIN, both when looking at overall data and when splitting the data by gender ($p>0.13$ for all comparisons, Wilcoxon signed-rank tests). In AMBIGUOUS, however, subjects increase their performance in Stage 3 relative to Stage 2, (8.75 in Stage 2 vs. 9.45 in Stage 3; $p<0.01$). This effect is significant for men ($p<0.01$) and for women ($p=0.03$).

Uncertain tournament environments may induce different performance levels in Stage 3 when comparing subjects across treatments (see panels c and d in Table 2 and the p -values reported there). Considering all subjects regardless of their competition entry choice (in panel c), we see that performance by men increases significantly in the presence of uncertainty or ambiguity about the number of winners, and that this increase is quite sizeable, namely 14.5% in UNCERTAIN and 15.4% in AMBIGUOUS. The corresponding differences for women are insignificant. Disaggregating this analysis by Stage 3 choice (see panel d of Table 2), we see that men who enter competition have a performance of 9.08 in CERTAIN, which is significantly lower than the performance of 10.61 in UNCERTAIN and 10.78 in AMBIGUOUS. Contrary to this, women who select the tournament have on average a *lower* performance in UNCERTAIN and in AMBIGUOUS than in CERTAIN, although treatment differences are not significant. Moreover, as expected, there is no significant difference in performance across treatments for those subjects (male, female, or pooled) who chose the piece-rate payment scheme in Stage 3.¹⁴

¹³ We confirm with a series of Kruskal-Wallis tests that there is no difference in performance across treatments for any of the two genders in Stage 1 or Stage 2 ($p>0.52$ for all comparisons), which indicates that randomization has been successful.

¹⁴ For completeness we note here that subjects who selected competition had a higher performance overall than those who opted for the piece rate in Stage 3 ($p<0.01$, Mann-Whitney test). Aggregating over all three treatments,

Table 2. Performance in Stages 1, 2 and 3, by Treatment and Gender

	Panel a: Stage 1			Panel b: Stage 2		
	Pooled	Men	Women	Pooled	Men	Women
CERTAIN	8.02 (0.52)	7.96 (0.83)	8.08 (0.64)	8.73 (0.57)	8.58 (0.92)	8.88 (0.71)
UNCERTAIN	7.92 (0.37)	7.98 (0.6)	7.85 (0.42)	8.89 (0.37)	9.08 (0.58)	8.69 (0.45)
AMBIGUOUS	7.99 (0.32)	8.38 (0.46)	7.60 (0.43)	8.75 (0.29)	9.04 (0.41)	8.46 (0.41)
Overall	7.97 (0.22)	8.13 (0.34)	7.80 (0.27)	8.80 (0.22)	8.97 (0.34)	8.63 (0.28)
CERTAIN vs. UNCERTAIN	$p=0.949$	$p=0.792$	$p=0.782$	$p=0.540$	$p=0.335$	$p=0.805$
CERTAIN vs. AMBIGUOUS	$p=0.745$	$p=0.313$	$p=0.528$	$p=0.574$	$p=0.272$	$p=0.718$

Panel c: Stage 3 (all subjects)

	Pooled	Men	Women
CERTAIN	9.00 (0.58)	8.54 (0.99)	9.46 (0.61)
UNCERTAIN	9.27 (0.36)	9.77 (0.53)	8.77 (0.47)
AMBIGUOUS	9.45 (0.30)	9.85 (0.42)	9.04 (0.43)
OVERALL	9.29 (0.22)	9.56 (0.34)	9.02 (0.28)
CERTAIN vs. UNCERTAIN	$p=0.393$	$p=0.051$	$p=0.337$
CERTAIN vs. AMBIGUOUS	$p=0.150$	$p=0.021$	$p=0.589$

this is true for men ($p<0.01$) as well as for women ($p=0.02$). Looking at single treatments, the difference is insignificant in CERTAIN, but significant in UNCERTAIN and AMBIGUOUS ($p<0.01$ in the latter two treatments).

Panel d: Stage 3 (conditional on choice)

	Selected Tournament			Selected Piece-Rate		
	Pooled	Men	Women	Pooled	Men	Women
CERTAIN	9.78 (1.10)	9.08 (1.54)	11.17 (1.19)	8.53 (0.66)	8.00 (1.31)	8.89 (0.68)
UNCERTAIN	10.45 (0.57)	10.61 (0.73)	10.15 (0.89)	8.38 (0.42)	8.6 (0.69)	8.26 (0.53)
AMBIGUOUS	10.26 (0.43)	10.78 (0.5)	9.13 (0.79)	8.67 (0.39)	8.00 (0.55)	9.00 (0.51)
Overall	10.25 (0.34)	10.43 (0.44)	9.88 (0.53)	8.52 (0.27)	8.25 (0.46)	8.67 (0.32)
CERTAIN vs. UNCERTAIN	$p=0.279$	$p=0.078$	$p=0.506$	$p=0.937$	$p=0.389$	$p=0.364$
CERTAIN vs. AMBIGUOUS	$p=0.275$	$p=0.036$	$p=0.224$	$p=0.565$	$p=0.498$	$p=0.929$

Standard errors in parentheses. All p values refer to Mann-Whitney U tests, comparing performance across treatments

In Table 3 we present a regression analysis of performance in Stages 1, 2 and 3. The right-hand side variables include the treatment dummies, gender (in the pooled specifications 1-3), the risk and ambiguity measures, and the economics dummy. In column (3) we also add interactions between gender and treatment. In columns (3)-(5), where the dependent variable is Stage 3 performance, the list of explanatory variables additionally includes Stage 2 performance, as a measure of ability. This allows us to control for the possibility that some of the results reported so far may be driven by imbalances in ability across treatment due to chance.

The first two columns of Table 3 confirm the results of the nonparametric tests on the absence of gender or treatment differences in Stage 1 and Stage 2 performance. The regression results also confirm the previous analysis on Stage 3 performance based on Table 2. In the case of women, the effects of uncertainty and ambiguity on performance are negative, but insignificant. The significant interaction terms between gender and treatment dummies in column (3) reveal that uncertainty and ambiguity affect men and women differently. Indeed, the effect of uncertainty and ambiguity on male performance is significantly positive. This effect is captured by the treatment dummies in column (4) that only uses data from male subjects, and by the joint coefficients *Uncertain + Uncertain x Male* ($p=0.069$) and *Ambiguous*

+ *Ambiguous x Male* ($p=0.034$) in specification (3). The above analysis allows us to state our second result.

Result 2: Introducing uncertainty or ambiguity into the tournament setting increases the performance of men in Stage 3 significantly (compared to the baseline treatment CERTAIN), but there is no such effect on women's performance. Hence, gender differences in performance shift strongly in favor of male participants when the tournament involves uncertainty or ambiguity.

The precise reason why men – but not women – increase their performance in the presence of uncertainty and ambiguity cannot be identified in the framework of our experiment. It is possible that the additional uncertainty in UNCERTAIN and AMBIGUOUS compared to CERTAIN is perceived by participants as a more challenging environment, or as an environment with stronger competitive pressure, leading men to unlock their full potential. Existing evidence showing that men often perform better than women in response to competitive pressure (Gneezy et al., 2003; Iriberry and Rey-Biel, 2018) supports such a possible explanation.

Table 3. Regression Estimates for Performance

	(1)	(2)	(3)	(4)	(4)
	Pooled	Pooled	Pooled	Men	Women
Dependent variable:	Performance	Performance	Performance	Performance	Performance
	Stage 1	Stage 2	Stage 3	Stage 3	Stage 3
<i>Male</i>	-0.425 (1.006)	-0.191 (1.010)	-0.968* (0.585)		
<i>Uncertain</i>	-0.129 (0.865)	-0.017 (0.868)	-0.716 (0.503)	0.973* (0.570)	-0.693 (0.453)
<i>Ambiguous</i>	0.475 (0.862)	-0.333 (0.865)	-0.317 (0.501)	1.121** (0.567)	-0.282 (0.452)
<i>Correct2</i>			0.804*** (0.039)	0.793*** (0.056)	0.822*** (0.056)
<i>Uncertain x Male</i>	0.499 (1.237)	0.605 (1.241)	1.641** (0.719)		
<i>Ambiguous x Male</i>	1.191 (1.233)	0.840 (1.238)	1.398* (0.717)		
<i>Risk Measure</i>	0.302 (1.360)	0.119 (1.365)	0.907 (0.790)	1.239 (1.250)	0.568 (1.023)
<i>Ambiguity Measure</i>	-0.515 (1.527)	0.374 (1.533)	-0.123 (0.888)	-0.028 (1.582)	-0.138 (1.029)
<i>Economics</i>	0.655 (0.475)	0.600 (0.477)	0.801*** (0.277)	0.902** (0.421)	0.654* (0.368)
Constant	7.845*** (0.991)	8.562*** (0.995)	1.818*** (0.665)	0.715 (0.966)	1.859** (0.802)
Observations	231	231	231	115	116

Notes: Robust standard errors in (1) and (2), following Breusch-Pagan tests that reject the null of homoscedastic standard errors ($p=0.10$, $p=0.01$, respectively). Null hypothesis of homoscedastic standard errors not rejected in (3), (4) and (5) ($p=0.60$, $p=0.40$, $p=0.98$, respectively, Breusch-Pagan tests).

3.3 Gender composition of tournament winners

An important question in the context of gender and competition is the gender distribution of tournament winners. At the end of the day, do uncertainty and ambiguity affect the share of men and women who win the tournament in Stage 3 compared to the control treatment? The data analysis so far gives rise to the conjecture that the fraction of men winning the tournament should increase in UNCERTAIN and AMBIGUOUS due to two reasons: First, the large gender gap for those who rank in the top half (documented in Figure 2) is particularly worrisome in

terms of securing a gender-balanced pool of tournament winners, since individuals with better performances in Stage 2 are also most likely to win the tournament in Stage 3. Moreover, Result 2 states that men increase their performance under uncertainty and ambiguity, while women do not.

Indeed, this conjecture is borne out in the data: the two treatments with uncertainty and ambiguity are characterized by a more unbalanced pool of tournament winners in Stage 3 compared to the control treatment. Table 4 reports the proportion of men and women within the pool of tournament winners in Stage 3, broken down by treatment. Among winners, the proportion of male subjects is statistically equal to that of women in CERTAIN (58.33%, $p=0.56$; one-sample test of proportions), but it is significantly different from 50% in UNCERTAIN (72%, $p=0.03$) and in AMBIGUOUS (69.44%, $p=0.02$).¹⁵ This leads to our third, and final result.

Result 3: The gender composition of winners gets more unbalanced in UNCERTAIN and AMBIGUOUS. This is a consequence of men improving their performance (while women do not) and the gender gap in the willingness to compete increasing in absolute terms in these two treatments. Hence, under more realistic conditions (with uncertainty and ambiguity), tournaments become skewed towards more male winners.

Table 4. Proportion of each gender among tournament winners in Stage 3

	Men	Women	P value
CERTAIN	0.583	0.417	0.564
UNCERTAIN	0.720	0.280	0.028
AMBIGUOUS	0.694	0.306	0.020
OVERALL	0.685	0.315	0.001

P values refer to one-sample tests of proportions, testing whether the proportion of each gender among tournament winners is equal to 0.5

¹⁵ For completeness we note that in Stage 2 men win the tournament 58.3% of the time, which is not significantly different from 50% ($p=0.20$, one-sample test of proportions).

4 Discussion and Conclusion

Labor markets entail a considerable degree of uncertainty when workers compete for scarce jobs and higher pay. In many cases, the conditions of a tournament may not be fully known to competitors. While in some cases they may have accurate guesses about the potential number of winners and the likelihood of each number, in others the exact conditions may be ambiguous with unknown probabilities. Since the labor market outcomes of men and women are still markedly different (Blau and Kahn, 2017), an ever growing body of literature has focused on gender differences in competitive behavior in order to explain (at least partly) the different labor market outcomes. This literature has emphasized potential gender differences both in competitive performance (starting with Gneezy et al., 2003) and in the willingness to compete (initiated by Niederle and Vesterlund, 2007), often finding that women shy away more often from competition and perform worse under competitive pressure than men. Interestingly, this literature has concentrated on settings where the tournament environment is perfectly known to competitors, meaning that they have perfect information about the number of competitors and the number of winners. Only a few field experiments (Flory et al., 2015; Gee, 2018) have had situations in which the competitors lacked perfect knowledge. However, contrary to our study, these field experiments did not measure *both* competitive performance and willingness to compete. Arguably, both dimensions are important when it comes to the ultimate question of which gender is more successful in tournaments, which is determined by a combination of tournament entry and performance in the tournament. Our experimental design has been able to study both dimensions in a unified framework.

We have found that men significantly increase their performance when the number of tournament winners becomes uncertain or ambiguous, while the performance of women decreases on average. Uncertainty and ambiguity cause both men and women to increase their average likelihood to enter competition, with the effect being larger for men and significant in the case of the ambiguous tournament. On average, the gender gap in the willingness to compete increases with uncertainty and ambiguity, but not significantly so in comparison to the control treatment with certainty. In combination, the effects of uncertainty and ambiguity on performance and willingness to compete lead to the important finding that men win the tournament significantly more often than women whenever either uncertainty or ambiguity is involved. Seen from this perspective, the previous experimental studies with perfect information and certainty about the number of competitors and winners might have measured only a lower bound for the gender differences in competitive behavior and success under competitive environments.

Our results imply that competitive behavior of men and women appears to depend on the level of information available to potential competitors. The gender composition of winners (determined by tournament entry and performance in the competition of those who select into the tournament) seems to be better balanced when tournament conditions are perfectly known than when there is uncertainty or ambiguity involved. This suggests that the management of companies can promote a gender balance through information given to competitors, in particular by making the tournament conditions more transparent. This may even reduce costs of screening candidates as long as the best qualified persons still choose to compete when more information is disclosed, as has been the case in our experiment. Such an approach to promote a gender balance via changing the available information for competitors might be a non-expensive and much less invasive alternative to institutionalized interventions such as affirmation action programs. A better gender balance may even support the longevity of companies (Weber and Zulehner, 2009) and it may also provide more female role models for future cohorts that compete on labor markets.

References

- Ahn, D., Choi, S., Gale, D., and Kariv, S. (2014). Estimating ambiguity aversion in a portfolio choice experiment. *Quantitative Economics*, 5(2), 195-223.
- Almås, I., Cappelen, A. W., Salvanes, K. G., Sørensen, E. Ø., and Tungodden, B. (2016). Willingness to compete: Family matters. *Management Science*, 62(8), 2149-2162.
- Angelov, N., Johansson, P., and Lindahl, E. (2016). Parenthood and the gender gap in pay. *Journal of Labor Economics*, 34(3), 545-579.
- Antonovics, K., Arcidiacono, P., and Walsh, R. (2009). The effects of gender interactions in the lab and in the field. *Review of Economics and Statistics*, 91(1), 152-162.
- Balafoutas, L. and Sutter, M. (2012). Affirmative action policies promote women and do not harm efficiency in the laboratory. *Science*, 335(6068), 579-582.
- Banerjee, R., Datta Gupta, N., and Villeval, M. C. (2018). The Spillover effects of affirmative action on competitiveness and unethical behavior. *European Economic Review*, 101, 567-604.
- Blau, F. and Kahn, L. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55, 789-865.
- Bock, O., Baetge, I., and Nicklisch, A. (2014). Hroot – Hamburg registration and organization online tool. *European Economic Review*, 71, 117-120
- Bognanno, M. (2001). Corporate tournaments. *Labour Economics*, 19(2), 290-315.
- Booth, A. and Nolen, P. (2012). Choosing to compete: How different are girls and boys?. *Journal of Economic Behavior & Organization*, 81(2), 542-555.
- Borghans, L., Heckman, J. Golsteyn, B., and Meijers, H. (2009). Gender difference in risk aversion and ambiguity aversion. *Journal of the European Economic Association*, 7(2-3), 649-658.
- Brandts, J., Groenert, V., and Rott, C. (2015). The impact of advice on women's and men's selection into competition. *Management Science*, 61(5), 1018-1035.
- Buser, T., Dreber, A., and Mollerstrom, J. (2017). The impact of stress on tournament entry. *Experimental Economics*, 20(2), 506-530.
- Buser, T., Niederle, M., and Oosterbeek, H. (2014). Gender, competitiveness and career choices. *Quarterly Journal of Economics*, 129(3), 1409-1447.
- Calsamiglia, C., Franke, J., and Rey-Biel, P. (2013). The incentive effects of affirmative action in a real-effort tournament. *Journal of Public Economics*, 98, 15-31.

- Card, D., Cardoso, A., and Kline P. (2016). Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on relative pay of women. *Quarterly Journal of Economics*, 131(2), 633-686.
- Charness, G. and Gneezy, U. (2012). Gender differences in risk aversion and competition. *Journal of Economic Behavior & Organization*, 83(2), 50-58.
- Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47(2), 448-474.
- Dargnies, M.-P. (2012). Men too sometimes shy away from competition: The case of team competition. *Management Science*, 58, 1982-2000.
- Datta Gupta, N., Poulsen, A., and Villeval, M.C. (2013). Gender matching and competitiveness: Experimental evidence. *Economic Inquiry*, 51(1), pp.816-835.
- Dimmock, S. G., Kouwenberg, R., Mitchell, O. S., and Peijnenburg, K. (2015). Estimating ambiguity preferences and perceptions in multiple prior models: Evidence from the field. *Journal of Risk and Uncertainty*, 51(3), 219-244.
- Dreber, A., von Essen, E., and Ranehill, E. (2011). Outrunning the gender gap—boys and girls compete equally. *Experimental Economics*, 14(4), 567-582.
- Dreber, A., von Essen, E., and Ranehill, E. (2014). Gender and competition in adolescence: Task matters. *Experimental Economics*, 17(1), 154-172.
- Dutcher, E.G., Balafoutas, L., Lindner, F., Ryvkin, D. and Sutter, M. (2015). Strive to be first or avoid being last: An experiment on relative performance incentives. *Games and Economic Behavior*, 94, 39-56.
- Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, 643-669.
- Eriksson, T. (1999). Executive compensation and tournament theory: Empirical tests on Danish data. *Journal of Labor Economics*, 17(2), 262-280.
- Ertac, S., and Szentes, B. (2011). The effect of information on gender differences in competitiveness: Experimental evidence. Koc University-TÜSIAD Economic Research Forum Working paper 1104.
- Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), 171-178.
- Flory, J., Leibbrandt, A., and List, J. (2015). Do competitive work places deter female workers? A large-scale natural field experiment on gender differences. *Review of Economic Studies*, 82(1), 122-155.

- Freeman, R.B. and Gelber, A.M. (2010). Prize structure and information in tournaments: Experimental evidence. *American Economic Journal: Applied Economics*, 2(1), 149-64.
- Gee, L. (2018). The more you know: Information effects on job application rates in a large field experiment. *Management Science*, forthcoming.
- Ghirardato, P., Maccheroni, F., and Marinacci, M. (2004). Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory*, 118 (2), 133–173.
- Gilboa, I., and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18 (2), 141–153.
- Gill, D. and Prowse, D. (2014). Gender differences and dynamics in competition: The role of luck. *Quantitative Economics*, 5(2), 351-376.
- Gillen, B., Snowberg, E., and Yariv, L. (2018). Experimenting with measurement error: Techniques with applications to the Caltech Cohort Study. *Journal of Political Economy*, forthcoming
- Gneezy, U., Niederle, M., and Rustichini, A. (2003). Performance in competitive environments: Gender differences. *Quarterly Journal of Economics*, 118(3), 1049-1074.
- Gneezy, U. and Rustichini, A. (2004). Gender and competition at a young age. *American Economic Review: Papers & Proceedings*, 94(2), 377-381.
- Halevy, Y. (2007), Ellsberg revisited: An experimental study. *Econometrica*, 75, 503-536.
- Harbring, C. and Irlenbusch, B. (2008). How many winners are good to have?: On tournaments with sabotage. *Journal of Economic Behavior & Organization*, 65(3-4), 682-702.
- Healy, A. and Pate, J. (2011). Can teams help to close the gender competition gap? *Economic Journal*, 121(555), 1192-1204.
- Iriberri, N., Rey-Biel, P. (2018). Competitive pressure widens the gender gap in performance: Evidence from a two-stage competition in mathematics. *Economic Journal*, forthcoming.
- Kölle, F. (2017). Affirmative action and team performance. *Journal of Economic Psychology*, 62, 50-62.
- Leibbrandt, A. and List, J. (2014). Do women avoid salary negotiations? Evidence from a large-scale natural field experiment. *Management Science*, 61(9), 2016-2024.
- Maggian, V., and Montinari, N. (2017). The spillover effects of gender quotas on dishonesty. *Economics Letters*, 159, 33-36.
- McGee, A., McGee, P., and Pan, J. (2015). Performance pay, competitiveness, and the gender wage gap: Evidence from the United States. *Economics Letters*, 128, 35-38.

- Müller, W. and Schotter, A. (2010). Workaholics and dropouts in organizations. *Journal of the European Economic Association*, 8(4), 717-743.
- Niederle, M. and Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics*, 122(3), 1067-1101.
- Niederle, M., Segal, C., and Vesterlund, L. (2013). How costly is diversity? Affirmative action in light of gender differences in competitiveness. *Management Science*, 59(1), 1-16.
- Orrison, A., Schotter, A. and Weigelt, K. (2004). Multiperson tournaments: An experimental examination. *Management Science*, 50(2), 268-279.
- Petrie, R. and Segal, C. (2015). Gender differences in competitiveness: The role of prizes. GMU Working Paper in Economics No. 14-47.
- Powell, M. and Ansic, D. (1997). Gender differences in risk behavior in financial decision making: an experimental analysis. *Journal of Economic Psychology*, 18(6), 605-628.
- Rigdon, M. (2016). An experimental investigation into gender differences in wage negotiations. Working Paper, Rutgers University.
- Saccardo, S., Pietrasz, A., and Gneezy, U. (2018). On the size of the gender difference in competitiveness. *Management Science*, 64(4):1541-1554.
- Schubert, R., Gysler, M., Brown, M., and Brachinger, H-W. (2000). Gender specific attitudes towards risk and ambiguity: An experimental investigation. Working Paper, ETH Zurich.
- Sutter, M. and Glätzle-Rützler, D. (2015). Gender differences in the willingness to compete emerge early in life and persist. *Management Science*, 61(10), 2339-2354.
- Sutter, M., Glätzle-Rützler, D., Balafoutas, L., and Czermak, S. (2016). Cancelling out early age gender differences in competition: An analysis of policy interventions. *Experimental Economics*, 19(2), 412-432.
- Sutter, M., Kocher, M., Glätzle-Rützler, D., and Trautmann, S. (2013). Impatience and uncertainty: experimental decisions predict adolescents' field behavior. *American Economic Review*, 103(1), 510-531.
- Trautmann, S., van de Kuilen, G. (2015), Ambiguity attitudes. In: Keren, G., Wu, G. (eds.), *The Wiley Blackwell Handbook of Judgment and Decision Making*. Wiley, London. 89-116.
- Villeval, M. C. (2012). Ready, steady, compete. *Science*, 335(6068), 544-545.
- Weber, A. and Zulehner, C. (2009). Female hires and the success of start-up firms. *American Economic Review* 100, 358-361.
- Weichselbaumer, D. and Winter-Ebmer, R. (2007). The effects of competition and equal treatment laws on gender wage differentials. *Economic Policy*, 22(50), 236-287.

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

Appendix (for online publication only):

Gender differences in their beliefs about relative performance.

Table A1 reports elicited beliefs about one's rank within a group in Stages 1 and 2. Men report significantly lower perceived ranks than women for their performance (i.e., they are more confident about having a good rank) in Stage 1 and in Stage 2 ($p < 0.01$ for both stages, χ^2 tests).

Table A1. Elicited beliefs about ranks in Stages 1 and 2 (percentages)

	Stage 1		Stage 2	
	Men	Women	Men	Women
Rank 1 (best)	37.5%	16.7%	38.3%	10.0%
Rank 2	34.2%	39.2%	39.2%	50.8%
Rank 3	22.5%	33.3%	15.8%	30.8%
Rank 4 (worst)	5.8%	10.8%	6.7%	8.3%

Fraction of subjects stating the belief that their performance ranks as 1st, 2nd, 3rd, or last within their group. Sums may not add up to 100 due to rounding.

Risk and ambiguity attitudes elicited in Stage 4.

Table A2 summarizes the risk and ambiguity measures for men and women elicited in Stage 4 of the experiment. We find that men are, on average, slightly risk seeking and have an average risk measure of $r=0.47$ ($p=0.01$, Wilcoxon sign-rank test, testing whether r is different from 0.5) while women are very close to risk neutrality ($r=0.51$, $p=0.94$ when testing whether r is different from 0.5). As a result, women are more risk averse than men on average ($p=0.09$). Subjects in our sample are significantly ambiguity averse, $a=0.07$ ($p<0.01$, sign-rank test when testing whether a is different from 0). However, there is no difference in ambiguity aversion between men and women ($p=0.44$). Moreover, the differences in risk and ambiguity across treatments are small and insignificant ($p=0.32$, $p=0.67$, Kruskal-Wallis tests for risk and ambiguity, respectively).

Table A2. Risk and Ambiguity Measures

	Overall			CERTAIN		
	Pooled	Men	Women	Pooled	Men	Women
Risk	0.49	0.47	0.51	0.51	0.53	0.48
	(0.01)	(0.02)	(0.02)	(0.03)	(0.05)	(0.03)
Ambiguity	0.06	0.05	0.07	0.07	0.05	0.09
	(0.01)	(0.01)	(0.01)	(0.03)	(0.05)	(0.03)
	UNCERTAIN			AMBIGUOUS		
	Pooled	Men	Women	Pooled	Men	Women
Risk	0.48	0.43	0.53	0.49	0.48	0.50
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)
Ambiguity	0.06	0.07	0.05	0.06	0.03	0.09
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)

Standard deviations in parentheses.

The risk measure ranges from 0 (most risk seeking) to 1 (most risk averse), with 0.5 indicating risk neutrality. The ambiguity measure has positive values for ambiguity aversion, zero for ambiguity neutrality and negative values for ambiguity loving.

Table A3. Regression Estimates for Entry into Competition, by treatment

	CERTAIN	UNCERTAIN	AMBIGUOUS
<i>Male</i>	-0.576 (0.407)	-0.608 (0.320)	-0.149 (0.431)
<i>Belief2</i>	-0.187** (0.091)	-0.245*** (0.079)	-0.42 (0.115)
<i>Correct2</i>	0.023 (0.021)	0.049** (0.022)	0.017 (0.022)
<i>Risk Measure</i>	-1.873** (0.810)	-1.701** (0.679)	-1.056** (0.491)
<i>Risk Measure x Male</i>	1.758* (1.023)	1.819** (0.872)	0.758 (0.850)
<i>Ambiguity Measure</i>	0.110 (0.891)	-1.329 (0.854)	-0.847* (0.485)
<i>Ambiguity Measure x Male</i>	0.459 (1.070)	1.395 (1.137)	1.772 (1.101)
Observations	46	93	92
<i>Risk Measure + Risk Measure x Male</i>	$p = 0.857$	$p = 0.824$	$p = 0.666$
<i>Ambiguity Measure + Ambiguity Measure x Male</i>	$p = 0.258$	$p = 0.929$	$p = 0.345$

Robust standard errors in parentheses. ***, **, * indicates significance at the 1%, 5%, and 10% level, respectively.

Models are Probit with marginal effects reported. Dependent variable is choosing to enter competition in Stage 3.

Sample Instructions

Welcome to an experiment on decision making. We thank you for your participation! During the experiment, you and the other participants will be asked to make certain decisions. Your own decisions as well as the decisions of the other participants will determine your payment from the experiment, according to the rules that will be described in what follows. The experiment will be conducted on the computer. You make your decisions on the screen. All decisions and answers will remain confidential and anonymous.

The experiment consists of 4 stages. You will be paid from one of the first three stages (chosen randomly at the end of the experiment) plus any earnings from Stage 4. Your total earnings from the experiment will be the sum of your payments from each of the two paid stages (Stage 4 plus one stage from Stages 1, 2, or 3). You will receive instructions for each of the four stages, one after the other.

We will read the instructions aloud and then give you time for questions. Please do not hesitate to ask questions if anything is not clear. Please do not talk to each other during the experiment. If you have any questions, please raise your hand.

Stage 1:

Your task in Stage 1 is to solve correctly as many addition exercises as possible. To be more precise, you will have 4 minutes time in order to solve as many additions of five randomly selected two-digit numbers as possible, by entering the sum of the five numbers. You are not allowed to use calculators but you can write down the numbers and use the provided scrap paper for your calculations. You enter an answer by clicking with the mouse on the “Confirm” button. When you enter an answer, you immediately find out on the screen whether it was correct or not.

If Stage 1 is the stage selected for payment (among Stages 1-3), then you will receive **€1.00 for each correct answer** that you entered within the 4 minutes. Your payment is not reduced when you enter a wrong answer.

Directly before the start of this stage you will be given one minute in order to familiarize yourselves with the screen: During this time you can solve addition exercises, which do not count for the experiment. Afterwards, Stage 1 will begin.

Stage 2:

As in stage 1, you will have 4 minutes time in order to solve correctly as many addition exercises as possible. However, your payment in this stage depends on your performance relative to the performance of a group of participants.

Group composition: Each group consists of **4 participants**, 2 of whom are men and 2 are women. Groups are randomly formed and each participant stays in the **same group** until the end of the experiment.

If Stage 2 is the stage selected for payment (amongst Stages 1-3), then your payment depends on how many additions you have solved correctly in comparison with the other three participants in your group. The group member who has entered the most correct answers is the winner of the tournament. The winner receives **€4.00 per correct answer** each, while the other three members **do not receive any payment**. In case of a tie, the ranking among the members with equal performances is determined randomly. You will not be informed about the outcome of the tournament until the end of the experiment.

Stage 3: Choice between Piece-rate payment and Tournament payment

As in Stages 1 and 2, you will have 4 minutes in order to solve correctly as many addition exercises as possible. However, you must now choose your preferred payment method for your performance in Stage 3. You can either choose a Piece-rate payment or the Tournament payment.

If you choose the **Piece-rate payment**, then you will receive **€1.00 per correct answer**.

If you choose the **Tournament payment**, then your performance in Stage 3 will be evaluated in comparison to the performance of the other three group members **in Stage 2**. As a reminder: That is the stage that you have just completed.

[IN TREATMENT CERTAIN:]

If you enter more correct answers than two of your group members did in Stage 2, then you will receive **€2.00 per correct answer**. In other words, only one member of your group can have a Stage 2-performance which is higher than your Stage 3-performance, otherwise you receive **no payment** for this stage.

[IN TREATMENTS UNCERTAIN AND AMBIGUOUS]:

The number of Stage-2 performances of your group members that you need to beat in order to be a winner in this Stage is uncertain. In particular:

[IN TREATMENT UNCERTAIN:]

Imagine there is a bag full of **blue**, **red**, or **green** marbles. There is an **equal number** of blue, red, and green marbles. Therefore, **each marble** has a **1/3 chance** of being chosen.

[IN TREATMENT AMBIGUOUS:]

Imagine there is a bag full of **blue**, **red**, or **green** marbles. **You do not know how many marbles of each color are in the bag.**

[IN TREATMENTS UNCERTAIN AND AMBIGUOUS]:

If a **blue** marble is drawn, then your performance must be better than **all three** of your group members' Stage 2 performances in order for you to win, and if you do win you will receive **€4.00 per correct answer**. In other words, no member of your group can have a Stage 2 performance which is higher than your Stage 3 performance; otherwise you receive **no payment** for this stage.

If a **red** marble is drawn, then your performance must be better than **two** of your group members' Stage 2 performances in order for you to win, and if you do win you will receive **€2.00 per correct answer**. In other words, only one member of your group can have a Stage 2 performance which is higher than your Stage 3 performance, otherwise you receive **no payment** for this stage.

If a **green** marble is drawn, then your performance must be better than **one** of your group members' Stage 2 performances in order for you to win, and if you do win you will receive **€1.33 per correct answer**. In other words, only two members of your group can have a stage 2 performance which is higher than your Stage 3 performance, otherwise you receive **no payment** for this stage.

Stage 4 [risk task first]

In Stage 4 you will be presented with two different tasks privately. The two tasks are similar but differ in some critical aspects. Your choice will affect your payment for Stage 4.

This stage consists of two tasks. In total, you will have to make 40 decisions, 20 in Task 1 and 20 in Task 2. Two of these decisions will be paid for real; one decision from Task 1 and one decision from Task 2. For each task you have to make 20 choices between a sure amount of money and drawing a ball from the virtual bag. By drawing from the bag you may win €6.

The virtual bag has orange and white balls. When you decide to draw a ball from the virtual bag, a ball will randomly be chosen. If the drawn ball is orange, you receive €6. If the drawn ball is white, you get nothing.

In each task you have 20 decisions where each decision is one row on your computer screen. For each decision you must choose if you prefer a sure amount of money or drawing a ball from the virtual bag. These 20 decisions will look as follows on your screen:

- | | | | | | |
|------|-------------------|--------------------------|-----------|--------------------------|--------------------|
| [1] | draw form the bag | <input type="checkbox"/> | or | <input type="checkbox"/> | 0.30 Euro for sure |
| [2] | draw form the bag | <input type="checkbox"/> | or | <input type="checkbox"/> | 0.60 Euro for sure |
| [3] | draw form the bag | <input type="checkbox"/> | or | <input type="checkbox"/> | 0.90 Euro for sure |
| [4] | draw form the bag | <input type="checkbox"/> | or | <input type="checkbox"/> | 1.20 Euro for sure |
| [5] | draw form the bag | <input type="checkbox"/> | or | <input type="checkbox"/> | 1.50 Euro for sure |
| [6] | draw form the bag | <input type="checkbox"/> | or | <input type="checkbox"/> | 1.80 Euro for sure |
| [7] | draw form the bag | <input type="checkbox"/> | or | <input type="checkbox"/> | 2.10 Euro for sure |
| [8] | draw form the bag | <input type="checkbox"/> | or | <input type="checkbox"/> | 2.40 Euro for sure |
| [9] | draw form the bag | <input type="checkbox"/> | or | <input type="checkbox"/> | 2.70 Euro for sure |
| [10] | draw form the bag | <input type="checkbox"/> | or | <input type="checkbox"/> | 3.00 Euro for sure |
| [11] | draw form the bag | <input type="checkbox"/> | or | <input type="checkbox"/> | 3.30 Euro for sure |
| [12] | draw form the bag | <input type="checkbox"/> | or | <input type="checkbox"/> | 3.60 Euro for sure |
| [13] | draw form the bag | <input type="checkbox"/> | or | <input type="checkbox"/> | 3.90 Euro for sure |
| [14] | draw form the bag | <input type="checkbox"/> | or | <input type="checkbox"/> | 4.20 Euro for sure |
| [15] | draw form the bag | <input type="checkbox"/> | or | <input type="checkbox"/> | 4.50 Euro for sure |
| [16] | draw form the bag | <input type="checkbox"/> | or | <input type="checkbox"/> | 4.80 Euro for sure |
| [17] | draw form the bag | <input type="checkbox"/> | or | <input type="checkbox"/> | 5.10 Euro for sure |
| [18] | draw form the bag | <input type="checkbox"/> | or | <input type="checkbox"/> | 5.40 Euro for sure |
| [19] | draw form the bag | <input type="checkbox"/> | or | <input type="checkbox"/> | 5.70 Euro for sure |
| [20] | draw form the bag | <input type="checkbox"/> | or | <input type="checkbox"/> | 6.00 Euro for sure |

For example, if line 7 is set as the payout-relevant line in the first task, you will receive € 6 if you have decided to draw the ball in this row and this ball is orange, but if the ball drawn is white then you receive € 0. If you have decided in this line for the secure amount (2.10 €), you will receive the respective safe amount indicated in line 7. You will find out the result from Stage 4 at the end of the experiment.

Task 1 *[risk]*

The virtual bag is full of 20 orange and white marbles. Half (10) of the marbles are orange and half (10) of the marbles are white. If you choose the virtual bag and an orange marble is chosen

you receive €6. If a white marble is chosen you receive nothing. For the 20 choices on your screen, please select if you prefer to draw from the bag or take the sure amount of money instead.

Task 2 [*ambiguity*]

Now there is a different virtual bag full of 20 orange and white marbles. You do not know the number of orange and white marbles in this case, but each distribution of white and orange marbles is equally likely. If you choose the virtual bag and an orange marble is chosen you receive €6. If a white marble is chosen you receive nothing. For each of the 20 choices on your screen, please select if you prefer to draw from the bag or take sure amount of money instead.