HAL
open science

Experimental estimates of men's and women's willingness to compete: Does the gender of the partner matter?

Seeun Jung, Radu Vranceanu

To cite this version:

Seeun Jung, Radu Vranceanu. Experimental estimates of men's and women's willingness to compete: Does the gender of the partner matter?. 2017. hal-01457226

HAL Id: hal-01457226

https://essec.hal.science/hal-01457226v1

Preprint submitted on 6 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

EXPERIMENTAL ESTIMATES OF MEN'S AND WOMEN'S WILLINGNESS TO COMPETE: DOES THE GENDER OF THE PARTNER MATTER?

RESEARCH CENTER
SEEUN JUNG, RADU VRANCEANU
ESSEC WORKING PAPER 1701

JANUARY 2017

Experimental estimates of men's and women's willingness to compete: Does the gender of the partner matter?

SeEun Jung* and Radu Vranceanu†

Abstract

In a classical experiment, Niederle and Vesterlund (2007) used the dichotomous choice of individuals between a piece rate and a tournament payment scheme as an indication of their propensity to compete. This paper reports results from a two person interaction of a similar type to analyze whether the preference for competition is dependent on the gender of the partner. It introduces a Becker–DeGroot–Marschak mechanism to elicit individual willingness to compete (WTC), defined as the amount of money that makes an individual indifferent between the two compensation schemes. Even when controlling for risk aversion, past performance and overconfidence, the male WTC is €3.30 larger than the female WTC. The WTC instrument allows for a more precise analysis of the impact of the partner's gender on the taste for competition.

Keywords: Willingness-to-compete, Experiments, Gender effect, BDM mechanism.

JEL Classification: C91, D03

1 Introduction

In recent years, gender differences in competitiveness have often been discussed as an argument to explain the persistent gender gap in wages, education and career choices in developed economies (Buser et al. (2014)). The assumption according to which on average men might present a higher propensity to compete than women was first tested by Niederle and Vesterlund (2007) and has been corroborated and extended through many subsequent field and experimental studies (see the surveys by Niederle and Vesterlund (2011) and Niederle (2016)).

*Inha University, Department of Economics, Incheon, South Korea. E-mail: jse@inha.ac.kr

†*Corresponding author.* ESSEC Business School and THEMA. PB 50105, 95021 Cergy, France. E-mail: vranceanu@essec.fr.

The main goal of this paper is to analyze whether the observed difference in the propensity to compete of men and women depends on the gender of the partner, in a setting as close as possible to the original paper by Niederle and Vesterlund (2011). In other words, do men (women) have the same propensity to compete whether the competitor has the same or the opposite gender?

In terms of performance, existing empirical evidence suggests that women tend to perform worse when they compete with men than when they compete with other women. For instance, Gneezy et al. (2003) compared performance in single-gender tournaments involving three men and three women, with performance in mixed-gender tournaments and found that men's performance is stronger in the mixed environment, while women's performance declines significantly in a mixed environment; women seem to react by reducing their effort when they compete with men, while the opposite happens for men. Iriberri and Rey-Biel (2012) showed that women under-perform compared to men in competitive environments only when the task is perceived as a male task but not when it is perceived as a female task. Gneezy and Rustichini (2004) analyzed the behavior of 10-year-old children running 40 meters, either separately or in pairs. They found that boys and girls improve their performance the most when competing in mixed groups; yet girls competing against girls slow down. Backus et al. (2016) analyzed the performance of expert chess players and observed that on average female players gain fewer matches than male players of same ability. This outcome is essentially driven by the gender composition of the match, with women performing much worse when paired with a man than with a woman of the same ability.

There is probably no need to present the details of the experimental design by Niederle and Vesterlund (2007) (hereafter referred to as NV) because over time their analysis has reached the status of a "classical" experiment. The experiment comprises a real effort task, consisting in adding successive blocks of five 2-digit numbers during five minutes. In a first stage, participants execute the task and are paid of \$0.50 per correct answer, what is referred to as the piece rate (PR) compensation. Then participants are assigned to mixed-gender groups of four persons. In a second stage, they execute the task, but only the best performer is paid \$2.00 per correct answer, which is referred to as the tournament (TN) compensation. In a third stage, participants must choose between the piece rate and the tournament compensation scheme before executing the task. An original and powerful idea consisted in determining the winner in stage three with respect to the performance in the stage two tournament. The advantage of this method is to ensure that the decision to enter the tournament relies only on one's beliefs about his own performance and preference to compete, and not on his beliefs about the entry decision of his partner (Niederle et al. (2013)).[1] As a fundamental

[1]Furthermore, one participant's decision does not impose any externalities on his partner.

result, they showed that 73% of the male subjects chose the tournament, yet only 35% of the female subjects did so, although there was no performance gap between the two genders (including when conditioned by the choice of the compensation scheme). Furthermore, the authors showed that the gender gap does not vanish if controlling for overconfidence or risk aversion, a residual to be explained by the male and female subjects' different attitudes toward competition.[2]

One idiosyncratic feature of the NV design pertains to the gender composition of the groups and the gender information structure. Because the authors wanted to ensure that their results hold in a mixed-gender environment, groups were composed of two men and two women seated in rows. Participants were informed that they will be grouped with the other people in their rows. As Niederle and Vesterlund (2011) acknowledge (p.603), "although gender was never mentioned during the experiment, individuals could see their competitors and determine the gender composition of the group". Because groups comprised two men and two women, each woman competed against two men and one woman; each man competed against one man and two women, *and they knew it*. Thus, men and women subjects faced a different environment with respect to the genders of their adversaries. If individuals of a given gender do not decide on entering competition depending on the gender of the partners, then this design will introduce no bias. In the opposite case, what has been interpreted as "women shy away from competition" should be restated as "women shy away from competition with (a majority of) men", which is not quite the same thing.

To the best of our knowledge, two studies – Geraldes (2016) and Datta Gupta et al. (2013) – have directly addressed the question of whether the gender of the partner has an impact on the individuals' propensity to compete, and reached different conclusions. In the two papers, subjects must chose between a competitive payment scheme based on a two-person tournament, and an uncompetitive payment scheme, based on a piece rate scheme. By contrast with NV, where subjects who choose the tournament will compete against the past performance of their partners, in these experiments, the subjects' payment in the tournament condition depends on the performance of their partner in that task. The experiments thus involve a much richer range of beliefs, about one's own performance, but also about the beliefs and choices of all the other persons in the room. The analysis by Geraldes (2016) uses the same real effort task as NV (adding numbers). He asked subjects to choose between a tournament against a randomly assigned partner and 1/2 chance to earn the piece rate compensation (or nothing). Then those who chose the tournament were matched with other

[2]For this purpose, they use an additional task, where participants have to choose between payments obtained in past PR and TN tasks without executing the task, thus "sterilizing" the effect of the competition thrill. The gender effect persists even if a dummy variable representative of this "inactive" choice is brought into the decision model of the "active" choice.

subjects who had also chosen the tournament. There are three treatments: in the "replication" treatment, subjects in mixed gender sessions are attributed to pairs at random; in the "choice of gender" treatment, subjects can choose the gender of the partner, and in the last treatment, the task is run in all-women sessions. In his data, women tend to avoid competition in mixed group sessions (as revealed by NV), but not in the two other treatments, what he interprets as a proof that women avoid competition with men, but not avoid competition in general. In particular, in the "choice of gender" condition, women choose women competitors, and enter competition more often. In the study by Datta Gupta et al. (2013), subjects are also asked to execute a real effort task consisting in solving mazes as in Gneezy et al. (2003). They can be paid either piece rate or according to tournament compensation involving two players. This time, subjects are informed about the gender of their partner in the tournament. Hence, the study will feature male-male, female-female and male-female pairs, allowing direct analysis of the impact of the partner's gender on the decision to compete in a tournament. Thus, if someone who chooses the tournament is paired with a person who chooses the piece rate, he/she will win "by default", as he/she has no competitor. Their main experiment (100 subjects) is based on transmission of gender information according the pseudonym method often used in psychological research (subjects receive a pseudonym that reflects the gender, and this pseudonym is communicate to partners); in this condition, the gender of the partner has no significant impact on the decision to compete. Only in the particular case (40 subjects) when gender information is overemphasized (by telling subject that the pseudonym is representative of the gender of the partner) do men tend to compete more against women than against other men. However, contrary to what was suggested by Geraldes (2016), no such a difference is observed for women.

Given that the evidence on gender interaction in the choice of competition has been inconclusive so far, additional tests are needed. The type of task and the related gender stereotypes (Iriberri and Rey-Biel (2012)), how saliently the gender information is conveyed (Datta Gupta et al. (2013)), the overall structure of the experiment including the control variables can have important consequences on the results. One important caveat pertains to the measure of "pro-competitive behavior". All the above-mentioned studies examine the individuals' dichotomous choice between a competitive alternative (a tournament compensation) and an uncompetitive alternative (a piece rate scheme), as representative of their desire to compete (and respectively to avoid competition). Such analyses say nothing about the "strength" of this preference for each person. One cannot rule out that this instrument is just not precise enough to detect the gender interaction effect.

Therefore, the second goal of our paper is methodological. We introduce an original slider task

to elicit individual willingness to compete (WTC). Basically, we use a Becker–DeGroot–Marschak (BDM) selection mechanism (Becker et al. (1964)) to prompt subjects to report truthfully the amount of money that makes them indifferent between the €0.50 piece rate and the €1.00 tournament. This amount of money should vary from one individual to another and express the strength of their preference for (or against) competition, in a much richer way than the plain yes/no choice of the tournament against the piece rate compensation. Our WTC elicitation method can be seen as an alternative to the "price-equivalent" method developed by Ifcher and Zarghamee (2016). In their experiment (hereafter referred to as IZ), subjects were also organized in groups of four and asked to execute the same real effort task as in NV.[3] In their original task 4, participants were asked to choose between a \$2.00 tournament payment, and a piece-rate payment p per correct answer, with p varying between \$0.00 and \$2.00 by increments of \$0.10. They must indicate their preferred choices in a Multiple Price List (MPL) with 21 rows. The computer will draw a line at random, and the individual will execute the task being paid according to his stated preferred choice for that line. The strength of each subject preference for competition is given by switching point between competition and piece rate, what authors define as the "piece-rate equivalent" of the \$2.00 tournament. The advantage of our method relies on the direct implementation of the BDM mechanism with its known and desirable strategy-proof properties. Furthermore, we feel that the WTC is easier to interpret as a direct measure of the preference to compete.

In general, our instructions replicate the experiment by NV; in particular, we kept the key principle according to which, when choosing the tournament, individuals compete against the past performance of their partner, thus eliminating any strategic or altruistic motives. To study "pure" gender interactions, we made groups of two individuals only (Datta Gupta et al. (2013); Buser (2016); Geraldes (2016)). The payoff structure is similar to NV too: in the piece rate treatment, each participant can earn €0.50 per correct answer; in the tournament treatment, the best performer on the team earns €1.00 per correct answer, and the other obtains nothing. All experimental sessions were mixed, but, within each session, we create both mixed gender pairs and homogenous gender pairs by random allocation (Datta Gupta et al. (2013)). We ensure that in tasks that involve interactions, participants are aware of the gender of their partner, using the method introduced by Jung and Vranceanu (2016, 2017). Tasks 1 to 3 are identical to NV. Task 4 serves to measure individual WTC.

We then study whether this WTC is subject to gender differences and impacted by the gender

[3]The same group gender composition and information structure was used by Ifcher and Zarghamee (2016). In their experiment, 24 out of 27 groups had the 2 men - 2 women composition.

of the partner. The analysis reveals that on average men are characterized by a preference for competition (their average WTC is 2.11) while women are characterized by competition aversion (their average WTC is -1.39). This gap persists even when controlling for performance, overconfidence and risk aversion. On average, men paired with men have a stronger WTC than man paired with women, and women paired with women have a lower aversion to competing than women paired with men, but these differences are not statistically significant. In our data, only overconfidence seems to be impacted by the gender of the partner, as male subjects are more overconfident when paired with women than with men.

The paper is organized as follows. The next section introduces the experimental design. Section 3 presents the results and the last section is our conclusion.

2 Experimental design

2.1 Subjects and sessions

The experiment was conducted at the ESSEC Business School (France) using students recruited from the pool of subjects of the Essec Experimental Lab.[4] Five sessions were organized at the ESSEC Experimental Lab with a total of 110 subjects as summarized in Table 1. Interaction is anonymous. They play the game on a computer screen and cannot establish eye contact with one another. Instructions (provided in Appendix B) and data collection were computerized; the program was developed using z-Tree (Fischbacher (2007)).[5]

An experimental session consists in subjects executing four main tasks (tasks 1 to 4), in which they add up sets of five 2-digit numbers for 5 minutes. In task 1, subjects obtain the piece rate pay of €0.50 per correct answer. In task-2, subjects are paired at random and learn the gender of the partner. They are paid according to a tournament scheme, the winner obtaining €1.00 per correct answer. In task 3, they must chose between piece-rate or tournament and then execute the task according to the preferred choice. In task 4, we elicit their WTC through an original slider task, then they execute their preferred option following a random draw by the computer. In tasks 2 to 4, subjects do not receive feedback on the performance of the partner until the very end of the session. After completing these main real effort tasks, subjects are asked to guess their rank

[4] As "Grande Ecole" students selected through a national competitive exam, this group is relatively homogenous in terms of computing and cognitive abilities, age and educational background. See the Lab website : http://behavioralresearchlab.essec.edu/

[5] We thank Delphine Dubart for the program development and implementation.

in task 2-tournament (can be first or second). They then complete an incentivized risk preference task, and complete a questionnaire to report on their risk tolerance and taste for competition.

For the purpose of this experiment, it was important to convey information about the gender of the partner without relaxing anonymity. We also sought to avoid the transmission of the gender information in a too salient way, which could have generated an undesired experimenter demand effect through an over-amplification of the gender stereotypes. The method used by Jung and Vranceanu (2016, 2017) seems to fulfil this requirement. At the very beginning of the experiment, immediately after the basic organization instructions, subjects must report their age, gender and education level. After the piece-rate task (task 1) and before the first competitive task begins (task 2), we inform them that they will execute the task paired with an anonymous subject, and we immediately release information about the gender, age, education level of the partner. However, in our student population, *only gender is a differentiating characteristic*. Thus, subjects are brought to know the gender of the partner, but have little reason to think that the experiment is first and foremost about gender. At the end of the experiment subjects must answer whether they remember the gender of the partner, and indicate it. One participant could not remember it, so we removed him from the analysis. The intensity of the gender information transmission in this method is probably weaker but close to the "standard" pseudonym method used by Datta Gupta et al. (2013). It is certainly less strong than the pseudonym plus gender statement in Datta Gupta et al. (2013) or the direct gender information transmission in Iriberri and Rey-Biel (2012).

All sessions were mixed, with approximately the same numbers of men and women in the room. In tasks 2 to 4, subjects were randomly allocated to pairs conditional on gender, to obtain Man-Man (MM), Woman-Woman (WW) and mixed (WM/MW) pairs (see Table 1). Subjects can participate only once in the experiment, which ensures the between-subjects configuration of the experiment.

Table 1: Summary of the Experimental Sessions

	Session 1	Session 2	Session 3	Session 4	Session 5	Total
Date	2016-10-04	2016-10-06	2016-11-08	2016-11-09	2016-11-14	
Nb participants	18	22	26	22	22	110
Nb Pairs MM	3	5	5	1	0	14
Nb Pairs WW	2	4	1	3	6	16
Nb Pairs MW	4	2	7	7	5	25
Nb Pairs	9	11	13	11	11	55

The final cash payment includes payment for one of the four main tasks selected at random at the end of the experiment, a participation amount of €5, in addition to the payment for the complementary tasks (risk taking, and rank guess). Experimental sessions last approximately 66 minutes; subjects received an average payment of €16.72.

We now present the details of each task.

2.2 Tasks 1 to 3

The first screen presents the organization of the experiment in four main tasks and some additional tasks and the payment system of the experiment.

As an original development, the second screen asks the participant to provide information on his/her age, education level and gender.

Task 1 perfectly replicates the instructions and steps of the NV experiment. Subjects are given five minutes to add chains of five 2-digit numbers displayed in the following way:

21	35	48	29	83	=??

For each such a problem, subjects had to fill in the sum in the last blank box. Subjects had pencils and scrap papers but were not allowed to use electronic devices to make the summations. Immediately after validating an answer, the computer indicates whether the answer is correct and recalls the total number of wrong/right answers. It then displays a new block of 2-digit numbers chosen at random by the computer. At the end of the 5 minutes, the task ends. Subjects are informed about the number of right answers in and the payment, then move to task 2.

For tasks 2 to 4, subjects are matched at random in groups of two. The composition of the group does not change from one task to another. Before task 2 begins, they are informed about the personal characteristics of their partner (thus, at that moment, they will learn the gender of the partner).

Task 2 is a basic tournament (or winner-takes-all) compensation scheme, where only the best performer on a team will earn €1.00 per correct answer, and the other partner will earn nothing. With two competitors in the tournament, a participant with a 1/2 chance of winning the tournament receives the same expected payoff as from the piece rate; this expectation payoff equivalence is similar to NV. In our set-up, 0.5 corresponds to the probability assigned to wining by a rational individual who has absolutely no clue about his relative ability with respect to the ability of his partner.

In task 3, participants must select which of the piece rate (€0.50) or tournament (€1.00 or €0.00) compensation schemes they want to be applied to their future performance. Like in NV

(and IZ), if they choose the tournament, their performance will be compared to the performance of their partner in task 2. Thus, the decision to compete reveals "pure" taste for competition, regardless of any strategic motives that would bring into the picture beliefs about changes in the partner's willingness to compete.

Notice that tasks 2 and 3 have a structure similar to tasks 2 and 3 in NV; they differ only in the number of competitors in a tournament (two vs. four in NV), and about the partners' gender information (explicit, vs. implicit in NV).

2.3 Task 4

Task 4 allows direct measurement of the individual's willingness to compete (WTC) according to a standard Becker–DeGroot–Marschak (BDM) mechanism, often used to elicit willingess-to-pay in consumer choices (Lusk and Shogren (2007)).[6]

Recall that in task 3, the subject indicated whether he/she prefers €0.50 PR or €1.00 TN, then executed the task according to the preferred option. In task 4, the participant is submitted to a similar choice: he/she can execute the real effort task and either be compensated with the €0.50 piece rate, or enter a tournament with possible payment of €1.00 per correct answer if he/she wins. Like in task 3, *his/her performance will be judged against the performance of his/her partner in task 2*. The main goal of task 4 is to determine the amount a that makes the subject indifferent between the two options. Let us denote an individual's valuation of the tournament compensation scheme by v_{TN} and his valuation of the piece rate payment by v_{PR}. We define a as $v_{TN} - v_{PR} = a$. If $a > 0$, the individual prefers the tournament scheme (*likes* competing), if $a < 0$, he/she prefers the piece-rate payment (*dislikes* competing). In other words, a is a direct measure of his/her WTC in this experiment.

Let us assume that we ask a participant to tell us his/her indifference amount, then let the computer choose between the two options at random with equal chances. Obviously, the individual has all the incentives to overstate the amount a. Several mechanisms, such as the second price auction, can induce truth–telling. In this paper, we use the Becker–DeGroot–Marschak mechanism. The individual is asked to state his indifference amount a, and then the computer draws at random an amount θ, with $\theta \in (-\bar{\theta}, \bar{\theta})$.[7] If $\theta > a$, the subject will execute the task and be paid the piece rate, plus the fixed amount θ; if $\theta < a$, the subject executes the task and is paid according to

[6] If we push the comparison to the anecdotal level, the choice between competing or not is not very different form asking an individual to choose between a sugar free drink and a similar but sugar based drink.

[7] Bohm et al. (1997) have shown that excessively high bounds or too narrow ones can alter the truth-telling property of the mechanism.

the tournament compensation. By dissociating the actual amount (θ) from the stated amount (a), revealing the true preference becomes a weakly dominating strategy. Thus, the BDM mechanism induces truth-telling regardless of individuals' risk preferences (Davis and Holt (1993)).

To implement this mechanism, we use a dynamic table (Figure 1) containing two blocks of symmetric options. In the upper pane, in each row 1 to 12 individuals have the choice between PR plus a fixed amount ($v_{PR} + \theta$) or the tournament (v_{TN}), with θ increasing on the upper rows. In the lower pane, rows 14 to 21, the subject can chose either the piece rate (v_{PR}) or the tournament plus a fixed amount θ, ($v_{TN} + \theta$). We allow θ to vary from €0.00 to €6.00 in increments of €0.50. Certainly, if the individual prefers $v_{PR} + \theta$ to v_{TN}, he/she should also prefer $v_{PR} + \theta'$ to v_{TN}, with $\theta' \geq \theta$. Symmetrically (in the lower pane), if the individual prefers $v_{TN} + \theta$ to v_{PR}, he should also prefer $v_{TN} + \theta'$ to v_{PR}, with $\theta' \geq \theta$. The middle row 13 corresponds to indifference between the TN and the PR schemes.

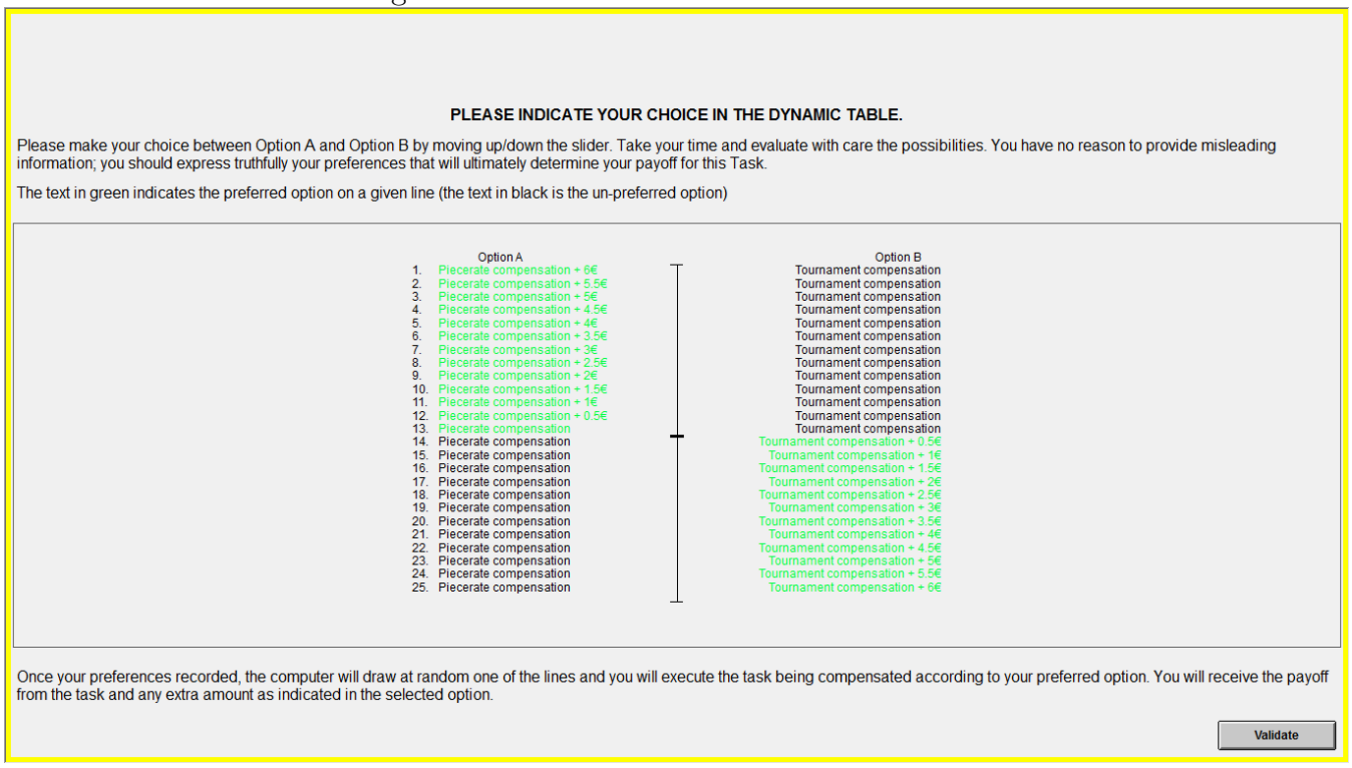
Individuals can select (in green) their preferred options by dragging a slider up or down. Initially the slider is set on the indifference row 13. By moving the slider up in the upper pane, the individual indicates that he/she has a stronger preference *for* the TN, as he requires an increasing additional compensation to play the experiment according to the PR. By moving the slider down in the lower pane, he/she indicates that he/she has a stronger preference *against* the TN, as he/she requires a higher amount to give up the PR and play the TN. The advantage of using a slider, instead of asking individuals to fill in the answer for each of the 25 rows, is that it rules out inconsistent choices where subjects can shift more than once.[8]

Once participants indicate their choice, the computer will draw one of the 25 lines at random, and the subjects will execute the task according to the choice indicated as preferred on that line. If the preferred choice contains a fixed amount $\theta > 0$, the individual receives it regardless of his performance (for instance, even if he/she loses in the tournament).

With 0 indicating the strict indifference between TN and PR, the table can elicit the subjects' WTC (expressed directly in EUR (€)). Positive numbers indicate a preference for competition, while negative numbers characterize competition aversion.

[8] Ifcher and Zarghamee (2016), who used a multiple price list, report that of 108 subjects, 8 presented an irrational reaction (shifted between TN and PR more than once). It is true that by "forcing" rationality, we might miss some of the natural errors humans can make, which can bring some interesting information to the researcher. However, many experiments drop these observations from the analysis.

Figure 1: The slider scale for WTC elicitation



2.4 Last steps

After completing the main tasks 1 to 4, subjects had to provide an incentivized guess about their rank in the task 2 tournament (as in NV and IZ). They could chose between [first, second] and were paid €1.00 for a correct answer.

Then, participants were submitted to an incentivized task aiming to grasp their attitudes towards risk. Many studies use the classical Holt and Laury (2002) test or its variants, in which subjects must select the preferred lottery in a list of two lotteries with varying means and standard deviations. If such a task can identify with precision an individual's degree of risk aversion, it relies on a relatively complex comparison of expected payoffs (Crosetto and Filippin (2016)). To avoid boredom at the end of our complex experiment, we resort to a new task, developed by Crosetto and Filippin (2013), the (Time) Bomb Risk Elicitation Task (BRET) to elicit individual risk tolerance.[9] At the beginning of the task, the computer displays 100 boxes, displayed in ten rows, each containing ten squared boxes. Each second, one more box is collected in the participant's basket (it changes color), from the left to the right and from the top to the bottom row. Boxes collected bring a payment of €0.20 per box. However, one box, with a random location, contains a "time bomb". If collected, this box sets all gains to zero. Subjects must decide at which point to stop collecting the boxes by pressing

[9] We thanks the authors for sharing the z-Tree program with the scholar community. https://paolocrosetto.wordpress.com/research/bret/

the "stop" button. Of course, the larger the number of boxes collected, the higher the potential gain, but also the higher the probability to collect the bomb and obtain nothing. The higher the number of boxes collected, the higher the subject's risk tolerance. As indicated by its authors, the BRET task requires minimal numeracy skills, avoids truncation of the data as subjects are free to choose any number between 0 and 100, allows precise estimation of both risk aversion and risk seeking, and is not affected by the degree of loss aversion. At the same time, the risk involved has a dynamic component that is not present in lottery choices and which might be attributed to a different dimension of risk taking.

A questionnaire at the end of the session elicits unincentivized measures of risk tolerance and taste for competition. We follow Dohmen et al. (2011) and asked subjects, "How do you see yourself ? Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks". The answer was on a scale from 0 (unwilling to take risks) to 10 (fully prepared to take risks). Authors found that this simple question is well correlated with incentivized lottery choices.[10] Using this risk measure in subsequent analyses will tell us something about using a lottery based measure of risk aversion as a complement to the BRET measure.

Finally, we also follow Buser (2016), who elicited pro-competitive behavior by asking subjects the simple question "How do you see yourself? Do you like to compete, or your rather prefer to avoid competition?", with answers on similar scale, where 0 stands for "I do not like to compete at all ", and 10 means "I really enjoy competing".

At the end of the session, subjects were informed about the results, and the composition and total amount of the final payment.

3 Results

The basic summary statistics are presented in the Appendix A (Tables A1, A2, and A3). One male subject who did not remember the gender of the partner was removed from the dataset. Three subjects (two men and one woman) showed inconsistent choices between task 3 and task 4 (preferred a different compensation scheme in task 4 compared to task 3) and were also removed from our analyses. The dataset thus contains 44 men and 47 women. In terms of pairs, there are 21 MM teams, 26 WW teams, and 44 mixed teams.

[10] It is also correlated with risk-taking in in a number of contexts including holding stocks, occupational choices and smoking.

3.1 Comparison with previous studies

Our performance results corroborate the findings by Niederle and Vesterlund (2007) and Ifcher and Zarghamee (2016) as we find no difference in unconditional performance between men and women in tasks 1 to 4 (Table 2).

Table 2: Comparison with other studies

	NV2007			ZI2016			This paper		
	Men	Women	p value	Men	Women	p value	Men	Women	p value
Task-1 performance	10.68	10.15	0.46	9.6	8.8	0.40	8.32	8.15	0.82
Task-2 performance	12.1	11.8	0.64	10.6	10.5	–	10.43	9.74	0.37
Task-3 performance	–	–	–	11.4	11.0	–	10.91	10.51	0.64
Task-4 performance	–	–	–	11.6	11.1	–	11.27	10.47	0.29
Competition chosen in Task-3	73%	35%		57%	34%		70%	30%	0.00
N	40	40		58	50		44	47	

The p values correspond to two-tailed t-tests.

Furthermore, 70% of the male subjects chose the tournament in task-3, compared to 30% of the women, another result that is quite similar to NV.

Following the steps of the previous studies, we analyze whether this substantial difference in choice can be explained by the differences in ability to perform the task. Table 3 presents the performance of men and women in task 1 and, in task 2, as well as the performance change between task 1 and task 2, contingent on their choice (PR or TN) in task 3 (similar tables are provided by NV and IZ).

Our results perfectly match the results by Niederle and Vesterlund (2007). There is no significant difference in the performance of women who chose the TN and those who chose the PR on all the three dimensions of performance. The individual ability as measured by performance in task 1 is similar for persons who chose PN or TN, for both the male and the female populations. The same claim holds for men, with the only exception, also detected in NV data, that in task 2, men who chose TN seem to have a marginally higher ability than those who chose PR. We must remark that in Ifcher and Zarghamee (2016) data, women who chose the TN have improved their performance more than those who chose the PR.

3.2 Overconfidence

An indicator that retained special attention in former studies is "over/underconfidence", defined as a positive/negative gap between the self-assessed and the actual rank in the task 2, tournament.

Table 3: Task-1 and -2 performance by gender and choice of task-3 payment scheme

Gender	Choice of payment scheme in task-3	Average performance			Observations
		Task-1	Task-2	Task-2 - Task-1	
		(1)	(2)	(3)	
Women	PR payment	7.91 (0.42)	9.30 (0.47)	1.39 (0.42)	33
	TN payment	8.71 (1.17)	10.79 (1.11)	2.07 (0.77)	14
	p value	0.42	0.15	0.41	
Men	PR payment	7.15 (0.68)	8.85 (0.59)	1.69 (0.52)	13
	TN payment	8.81 (0.58)	11.10 (0.79)	2.29 (0.41)	31
	p value	0.19	0.09	0.41	

Standard errors reported in parentheses. The p values correspond to two-tailed tests of equal performance for those who chose PR versus TN payment in Task 3

With only two players, overconfidence appears if one believes he/she is the first when he/she is second; underconfidence appears if one believes he/she is second when he/she is the first. Table 4 indicates the distribution of right/wrong perceptions regarding the rank in task 2 (tournament), depending on the gender of the subject. If the frequency of overconfident persons is the same in both the populations of men and women, women seem to be affected by underconfidence at a higher frequency than men (p-val: 0.08).

Table 4: Confidence

	Men	Women
Overconfidence	23% (0.06)	26% (0.06)
Right confidence	68% (0.07)	55% (0.07)
Underconfidence	9% (0.04)	19% (0.06)*

Standard errors reported in parentheses. The p values correspond to two-tailed tests of equal confidence for men and women.

However, Table 5 reports the same perception by gender, but this time depending on the gender of the partner. We observe that the consolidated figures in table 4 hide strong discrepancies in the men's beliefs: 26% of the men paired with women are overconfident, compared to 19% of men paired with men (p-val: 0.59). Furthermore, no man paired with a woman is underconfident, while 19% of men paired with men are underconfident (p-val: 0.03). The response of women is much more balanced, with women paired with men being less underconfident than women paired with

women.

Table 5: Confidence: gender composition

	Men		Women	
	Male Partner	Female Partner	Male Partner	Female Partner
Overconfidence	19% (0.09)	26% (0.09)	24% (0.10)	27% (0.09)
Right confidence	62% (0.11)	74% (0.09)	62% (0.11)	50% (0.10)
Underconfidence	19% (0.09)	0% (0.00)**	14% (0.08)	23% (0.08)

Standard errors reported in parentheses. The p values correspond to two-tailed tests of equal confidence for those with a male partner versus a female partner.

The perceived rank is thus an interesting indicator, conveying one individual's belief about his ability to succeed in a competition. A simple probit model (self rank: first - 1; second - 0) provides a better understanding of the determining factors of self-confidence.

There are two important indicator variables that capture the gender profile of individuals and teams - *Female subject* (FS, takes the value of 1 if subject is a woman, 0 otherwise) and *Female partner* (FP, takes the value of 1 if the partner is a woman, 0 otherwise). The "generic model" can be framed as:

$$y_i = X_i'\beta + \alpha_1 FS + \alpha_2 FP + FS \times FP + \epsilon_i \tag{1}$$

where y_i stands for the various *outcome variables* pertaining to an individual i, X is the vector of covariates other than the gender indicators and ϵ_i represents the standard error term. Note that in the presence of a significant interaction term $FS \times FP$, the coefficient α_1 indicates the marginal effect of the subject being a woman, conditional upon FP being zero (the partner is a man). Table 6 presents the output estimates of a probit model.

Columns 1 and 2 estimate the model (1) using as a risk control the incentivized risk task; columns 3 and 4 present estimates taking into account the self-assessed risk measure.

As shown by the significant coefficient α_2, *men paired with female partners tend to be overconfident (coefficient* $\alpha_2 > 0$); *women paired with women do not feature the same bias (*$\alpha_2 + \alpha_3 \approx 0$).

Furthermore, the task 2 performance is significant, indicating that a majority of subjects can correctly guess their ability to perform in this task.

The incentivized risk tolerance measure (BRET task) is not significant, while the self-reported risk tolerance measure is. This confirms that the two risk aversion measures may capture different dimensions of risk taking. In our sample, the correlation coefficient between the incentivized risk tolerance measure (BRET) and the self-reported risk tolerance is positive but small $\rho = 0.19$

Table 6: Self-rank: probit (first=1, second=0)

	(1)	(2)	(3)	(4)
Female subject	-0.425	0.184	-0.110	0.645
	(0.30)	(0.45)	(0.31)	(0.48)
Female partner		1.035**		1.308***
		(0.47)		(0.49)
Female subject × female partner		-1.233*		-1.424**
		(0.64)		(0.66)
Task-2 performance	0.290***	0.279***	0.333***	0.343***
	(0.07)	(0.07)	(0.08)	(0.08)
Improved Performance	-0.069	-0.041	-0.082	-0.062
(between task-1 and task-2)	(0.07)	(0.08)	(0.07)	(0.08)
Risk preference	0.009	0.008		
(incentivised bomb task)	(0.01)	(0.01)		
Risk preference			0.191***	0.217***
(self reported)			(0.07)	(0.07)
Constant	-2.512***	-2.928***	-3.662***	-4.584***
	(0.80)	(0.85)	(0.91)	(1.07)
N	91	91	91	91
chi2	25.886	31.066	33.062	40.733

Standard errors reported in parentheses. Significant levels are: * p < 0.1, ** p < 0.05, *** p < 0.01

($p < 0.10$).

3.3 Gender differences and interactions in task 3: Choosing to compete

The main aim of this study is to test for the presence of gender interactions in the propensity to compete in a tournament. In pursuit of this goal, Table 7 recalls frequencies of subject who chose tournament in task 3 by subject gender, and by subject gender conditional on gender of partner. This section thus takes the analysis of Niederle and Vesterlund (2007) one step further by controlling the decision to compete by the gender of the partner. As stated in the introduction, in the original experiment, male subjects had to compete against two women and one man, and vice versa for female subjects.

As we can see, there are no statistically significant differences in the proportion of men (respec-

Table 7: Gender interaction in Choosing to Compete (Task-3)

	Men	Women	Diff	P-value
	70% (0.07)	30% (0.07)	40% (0.09)***	0.000
N	44	47		
Male Partner	67% (0.11)	29% (0.10)	38% (0.15)**	0.013
N	21	21		
Female Partner	74% (0.09)	31% (0.09)	43% (0.13)***	0.002
N	23	26		
Diff	-7% (0.14)	-2% (0.14)		
P-value	0.609	0.873		

Standard errors reported in parentheses. The p values correspond to two-tailed tests of equal confidence for men versus woman and those with a male partner versus a female partner.

tively women) conditional on the gender of the partner, a result that corroborates the findings by Datta Gupta et al. (2013) from their weak information condition.[11] As already mentioned, in this experiment we also used a non-aggressive method to convey gender information. However, we are confident that individuals were aware of the gender of the partner, as all of the subjects in the database confirmed at the end of the session that they remembered the gender of the partner, and correctly indicate it.

Figures 2 and 3 depict the proportion of men who chose the tournament depending on whether the partner was a man or a woman, and the proportion of women who chose the tournament depending on whether their partner was a man or a woman, both depending on the task 2 performance (bottom 50%, top 50%) and self-rank (second, first).

Here again, we can see that a good performance and a high self-esteem seem to drive more men into competition, while these factors have a lesser impact on women's decision to compete. It appears that the effect of poor performance or poor self-confidence on the decision to avoid competition is amplified the case of women paired with men.

The most straightforward way to test for the impact of pair gender composition on the decision to compete in task 3 is by running a probit regression model with a new dependent variable, that takes the value of 1 if the individual opted for the tournament in task 3, and 0 if he/she opted for the piece rate.

[11] In their weak information condition, 34.4% of the women paired with a woman chose the tournament, and 33.3% of the women paired with a man chose the tournament. However, 55.5% of the men paired with a woman chose the tournament, and 62.5% of the men paired with a man chose the tournament.

Figure 2: Choosing to compete in task-3 and gender composition conditional on performance

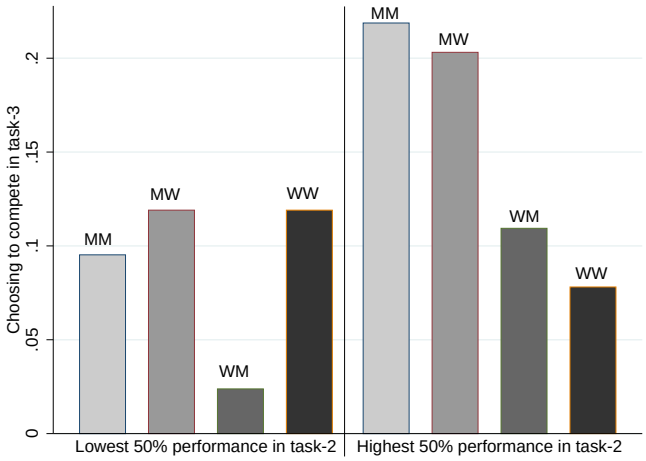


Figure 3: Choosing to compete in task-3 and gender composition conditional on self-rank

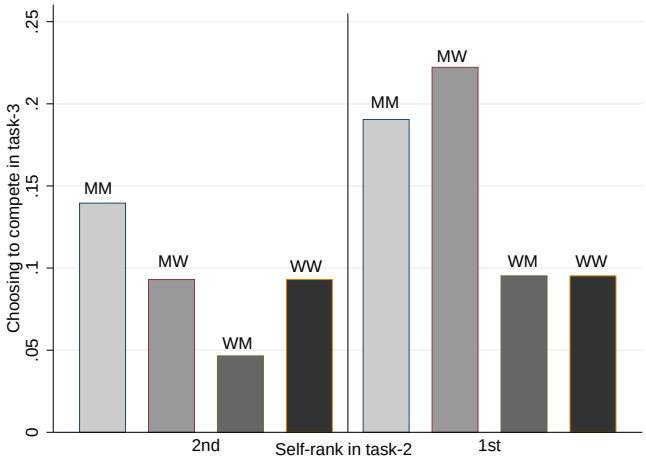


Table 8 presents the results including the incentivized risk measure as a covariate.[12]

Table 8: Choosing to compete in task-3: probit

	(1)	(2)	(3)
Female subject	-1.092***	-1.150**	-1.012**
	(0.30)	(0.46)	(0.44)
Female partner		0.006	0.254
		(0.45)	(0.43)
Female subject × female partner		0.104	-0.207
		(0.62)	(0.59)
Task-2 performance	0.042	0.042	0.091*
	(0.06)	(0.06)	(0.05)
Improved Performance	0.002	-0.000	-0.008
(between task-1 and task-2)	(0.07)	(0.07)	(0.07)
Task-2 self-rank	0.720**	0.725**	
(first=1, second=0)	(0.34)	(0.35)	
Risk preference	0.021	0.021	0.021*
(incentivised bomb task)	(0.01)	(0.01)	(0.01)
Constant	-1.158*	-1.162*	-1.312*
	(0.66)	(0.69)	(0.69)
N	91	91	91
chi2	26.864	26.933	22.534

Standard errors reported in parentheses. Significant levels are: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In line with the standard results by Niederle and Vesterlund (2007), self-rank and gender are key determining factors of the decision to compete; women avoid competition even when we control for self-confidence and risk aversion. The self-rank variable seems to absorb the effect of the risk tolerance. When the self-reported risk tolerance is used instead of the BRET task, the opposite happens: the risk measure becomes significant, and the self-rank significance vanishes.

Furthermore, the gender interactions terms are not statistically significant, confirming that the gender of the partner is not directly influencing the decision to enter into a tournament. When sup-

[12]In Appendix Table A4 we present the same regression with the alternative, self-reported tolerance to risk measure instead of the BRET measure.

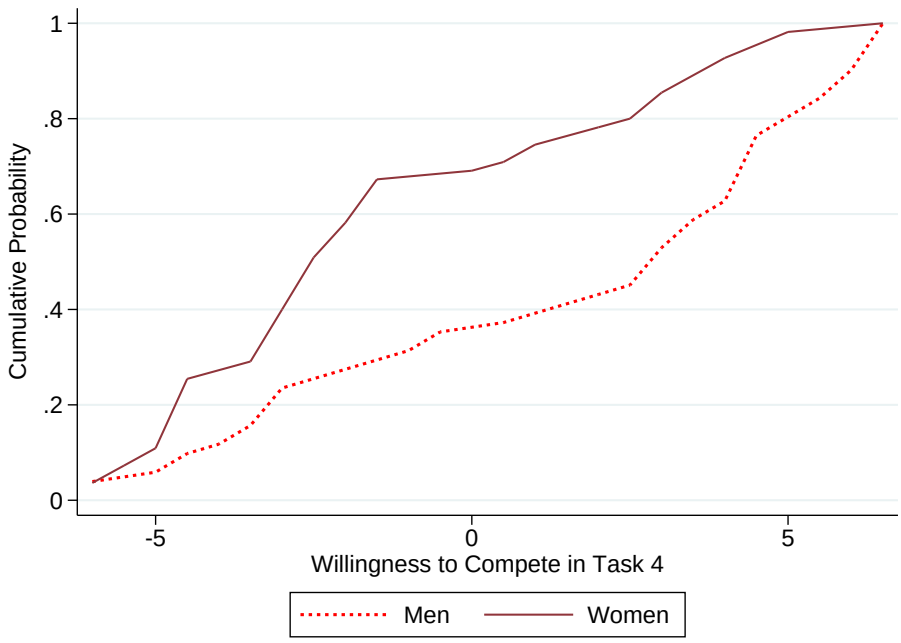
pressing self-rank from the analysis, the gender interaction terms remain not significant, indicating that the gender interaction is not absorbed by the overconfindence measure.

3.4 Gender differences and interactions in task 4: Willingness to compete

We now report the results from our task 4 aiming to measure the intensity of the preference between PR and TN of any subject. A positive number indicates the minimum amount required by a participant to refuse the tournament and execute the task according to the piece rate compensation. A negative number indicates the minimum amount required by a participant to give up the piece rate compensation and be compensated according to the tournament. High positive numbers reveal a high willingness to compete, negative numbers with a large absolute value indicate a strong aversion to competition, with zero as the neutral stance.

Figure 4 depicts the cumulative density distribution of WTC for each of the two subject populations of men and women (regardless of the gender of the partner). It can be seen that in these consolidated data, 70% of the men and 28% of the women report a positive WTC. Overall, the male WTC distribution stochastically dominates the female WTC distribution.

Figure 4: Cumulative density distribution of WTC by genders



As a robustness check, we verified that the WTC correlates with the self-reported measure of willingness to compete introduced by Buser (2016). The correlation coefficient is 0.479 (p-val<0.001), identical for male and female populations.

Table 9 presents the average WTC of the subjects, depending on their own gender and the gender of their partner.

Table 9: Gender interaction in WTC (Task-4)

	Men	Women	Diff	P-value
	2.11 (0.58)	-1.39 (0.50)	3.51 (0.76)***	0.000
N	44	47		
Male Partner	1.48 (0.81)	-1.74 (0.67)	3.21 (1.06)***	0.004
N	21	21		
Female Partner	2.70 (0.82)	-1.12 (0.73)	3.81 (1.09)***	0.001
N	23	26		
Diff	-1.22 (1.16)	-0.62 (1.01)		
P-value	0.149	0.541		

Standard errors reported in parentheses. Significant levels are: * p < 0.1, ** p < 0.05, *** p < 0.01

On average, male WTC is positive (€2.11), while female WTC is negative (€-1.39), showing that on average men are rather attracted by competition, while women dislike it. As already mentioned, our direct WTC measure is an alternative to the piece-rate equivalent method introduced by Ifcher and Zarghamee (2016). In their study's task 4, subjects had to chose between a 4-person tournament bringing $2.00 per correct answer and a piece rate varying from €0.00 to €2.00. After recording their choices on an MPL, the computer will draw one line, and they will execute the task accordingly. In their study, too, there are significant gender differences; for instance, on average men give up competition if the piece rate exceeds $0.94, while women would shift to piece rate if paid $0.66 or more. As a rough approximation, assuming that on average the number of right answers in their task 3 is 11.2 (Table 2), this correspond to a payoff difference of $3.13, which, as a magnitude, is quite close to our numbers (2.11+1.39=3.50).

However, a closer look into these WTC figures shows that gender average statistics can hide important differences with respect to the gender of the partner. In particular, men paired with women present a higher willingness to compete than men paired with men, and women paired with male partners have a higher aversion to competition (true, within our small sample, these differences are not statistically significant).

Figure 5 shows the distribution of WTCs depending on the performance in task 2. It turns out that best-performing (in task 2) men tend to compete more, but not women. There is a gender

Figure 5: WTC and gender composition conditional on performance

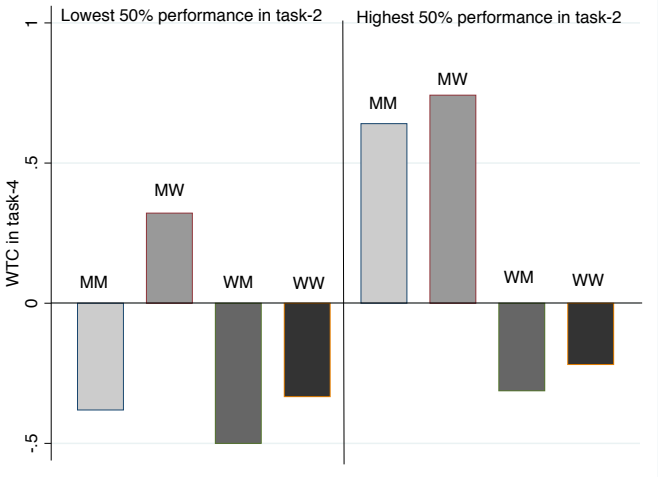
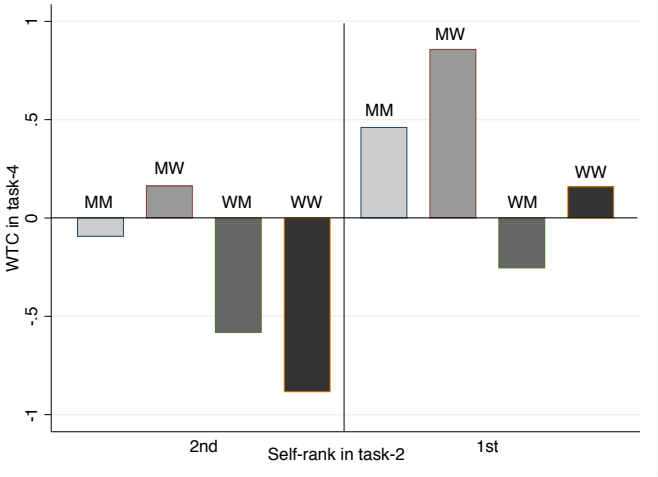


Figure 6: WTC and gender composition conditional on self-rank



group composition effect, but it does not seem to be very strong: poorly performing men matched with women prefer to compete, while they would shy away competition if paired with men.

Conditional on self-rank, male subjects present a higher WTC than women (Figure 6). Women paired with women will embrace competition if they have strong confidence in their strengths but not if paired with men.

Table 10 presents the output estimates of model (1), having individual WTC as the dependent variable. Dependent variables are the gender indicator variables, the performance measures, the self-rank in task 2 and the BRET risk tolerance measure. In Appendix Table A5 we present the same estimates using the self-reported risk tolerance instead of the BRET measure.

Like in NV and IZ, the self-rank and the (incentivized) risk tolerance seem to have a positive impact on the propensity to compete. Even when controlling for all these factors (performance, overconfidence and risk taking), the gender gap in WTC is large and statistically significant.

The estimates including the incentivized risk tolerance reveal a €3.30 WTC gap in favor of men

Table 10: Willingness to compete in task-4: OLS

	(1)	(2)	(3)
Female subject	-3.334***	-3.462***	-3.271***
	(0.76)	(1.15)	(1.16)
Female partner		0.640	1.162
		(1.14)	(1.11)
Female subject × female partner		0.198	-0.454
		(1.58)	(1.55)
Task-2 performance	0.155	0.160	0.247**
	(0.12)	(0.12)	(0.12)
Improved performance	-0.116	-0.129	-0.131
(between task-1 and task-2)	(0.17)	(0.18)	(0.18)
Task-2 self-rank	1.597*	1.510*	
(first=1, second=0)	(0.84)	(0.88)	
Risk preference	0.054*	0.054	0.057*
(incentivised bomb task)	(0.03)	(0.03)	(0.03)
Constant	-2.416	-2.744	-3.015*
	(1.62)	(1.72)	(1.73)
N	91	91	91
r2	0.293	0.302	0.277

Standard errors reported in parentheses. Significant levels are: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

(to be compared to €3.40 difference between the gender specific WTC averages). However, if the estimates include the self-reported risk tolerance instead of the BRET, the gap is reduced to about €1.93 according to model (2) in Table A5 (Appendix).

Furthermore, gender interactions do not impact directly the WTC. Even if suppressing the overconfidence term (which seems to depend on the gender of the partner), the gender interaction coefficients remain insignificant. This result holds if the alternative risk measure is brought into the picture (Appendix Table A5).[13]

As an upshot, if the mere comparison of the averages hints that the gender of the partner might have an impact on the decision to compete, despite a sensible instrument used to measure WTC,

[13]Only in model (3) in the Appendix, the partner gender coefficient is marginally significant (p-val<10%).

in our small sample, we could not find a statistically significant "partner gender" effect.

4 Conclusion

Many empirical studies have shown that individuals' performance of in various forms of competitive activities depends on the gender of the partner (Gneezy et al. (2003); Gneezy and Rustichini (2004); Iriberri and Rey-Biel (2012); Backus et al. (2016); Geraldes (2016)). However, evidence on whether the gender of the partner influences the decision to chose a tournament rather than an uncompetitive compensation scheme has been relatively inconclusive (Datta Gupta et al. (2013); Geraldes (2016)).

In this paper, we study whether a subject's choice of the tournament compensation depends on the gender of his/her partner in an experiment as close as possible to the classical experiment by Niederle and Vesterlund (2007). In addition to replicating their method involving a simple choice between the two payment systems, we also develop an original method, based on the Becker–DeGroot–Marschak mechanism, to infer an individual's willingness to compete. This method can be seen as an alternative to the price rate equivalent method introduced by Ifcher and Zarghamee (2016).

Our results can be added to a long list of experiments that corroborate the key findings by Niederle and Vesterlund (2007), namely that there is a strong gender effect in the choice of competition. Like in their classical paper, in our experiment 70% of the men and only 30% of the women chose the tournament in a binary choice, despite no significant differences in the conditional and unconditional performance of the two genders. As observed by Datta Gupta et al. (2013), and in contrast to Geraldes (2016), we find that the simple binary choice model (tournament vs. piece-rate) does not reveal any significant influence of the partner gender on the decision to compete. However, this negative result might be due to the low sensibility of the binary choice instrument. To test this hypothesis, our task 4 introduced a more sophisticated instrument able to capture individual willingness to compete on a continuous (-6, 6) scale. With a positive WTC of €2.11, on average male subjects have a preference for competition, and female subjects, with a negative WTC of €-1.39, present an aversion to competition. A €3.30 gender gap in WTC can still be observed after controlling for performance, overconfidence, and risk aversion in a regression model.

Overall, gender interactions do not appear to play a substantial role in this experiment. True, basic statistics show that men paired with women are the most competitive (WTC = 2.70), while men paired with men are less competitive (WTC = 1.48), a result also present in the "strong information" procedure of Datta Gupta et al. (2013). In line with the early result by Geraldes

(2016) women paired with men have a higher aversion to competition (WTC=-1.74) than women paired with women (WTC=-1.12), but none of these differences are statistical significant. No differences in WTC depending on the gender of the partner can be observed in a regression model that controls for performance, overconfidence and risk tolerance. Overconfidence seems to be the only variable affected by the gender of the partner, men being more overconfident when paired with women.

In the standard NV design, subjects of different gender are confronted with groups of three adversaries with different gender composition. The absence of a gender interaction effect should be seen as good news for these studies, as it removes doubts about the interpretation of the results in terms of pure gender differences. However, it is certainly too early to draw a definitive conclusion on this important topic. On the contrary, the question of gender interaction in choosing to compete is important enough to call for further replications and extensions of these analyses.

Acknowledgments. This research was funded by a grant provided by the ESSEC Alumni Foundation. The authors are grateful to Gorkem Celik for his contribution to the design of the slider task, to Aurélie Dariel for stimulating exchanges at the beginning of the project, and Delphine Dubart for her technical assistance.

References

Backus, P., M. Cubel, M. Guid, S. Sánchez-Pages, and E. L. Manas (2016). Gender, competition and performance:Evidence from real tournaments. Working Papers 2016-27, Institut d'Economia de Barcelona (IEB).

Becker, G. M., M. H. Degroot, and J. Marschak (1964). Measuring utility by a single-response sequential method. *Systems Research and Behavioral Science 9*(3), 226–232.

Bohm, P., J. Linden, and J. Sonnegard (1997). Eliciting Reservation Prices: Becker-DeGroot-Marschak Mechanisms vs. Markets. *Economic Journal 107*(443), 1079–89.

Buser, T. (2016). The Impact of Losing in a Competition on the Willingness to Seek Further Challenges. *Management Science* (4904).

Buser, T., M. Niederle, and H. Oosterbeek (2014). Gender, Competitiveness, and Career Choices. *The Quarterly Journal of Economics 129*(3), 1409–1447.

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

Crosetto, P. and A. Filippin (2016). A theoretical and experimental appraisal of four risk elicitation methods. *Experimental Economics 19*(3), 613–641.

Datta Gupta, N., A. Poulsen, and M. C. Villeval (2013). Gender Matching And Competitiveness: Experimental Evidence. *Economic Inquiry 51*(1), 816–835.

Davis, D. D. and C. A. Holt (1993). *Experimental Economics.* Princeton University Press.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011). Individual Risk Attitudes: Measurement, Determinants, And Behavioral Consequences. *Journal of the European Economic Association 9*(3), 522–550.

Fischbacher, U. (2007). z-Tree: Zurich Toolbox for Ready-made Economic Experiments. *Experimental Economics 10*(2), 171–178.

Geraldes, D. (2016). Women dislike competing against men. *Available online at grape.org.pl/wp-content/uploads/2016/08/Geraldes_paper.pdf*.

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

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

Holt, C. A. and S. K. Laury (2002). Risk aversion and incentive effects. *American Economic Review 92*(5), 1644–1655.

Ifcher, J. and H. Zarghamee (2016). Pricing competition: a new laboratory measure of gender differences in the willingness to compete. *Experimental Economics 19*(3), 642–662.

Iriberri, N. and P. Rey-Biel (2012). Let's (not) talk about sex: Gender awareness and stereotype-threat on performance under competition. *Barcelona GSE Working Paper Series*.

Jung, S. and R. Vranceanu (2016). Gender interaction in teams: Experimental evidence on performance and punishment behavior. *ESSEC working paper series* (WP1514).

Jung, S. and R. Vranceanu (2017). Experimental Evidence on Gender Interaction in Lying Behavior. *Revue Economique, Forthcoming*.

Lusk, J. L. and H. F. Shogren (2007). *Experimental Auctions*. Cambridge University Press.

Niederle, M. (2016). *Gender in Handbook of Experimental Economics*. Princeton University Press, Second Edition.

Niederle, M., C. Segal, and L. Vesterlund (2013). How Costly Is Diversity? Affirmative Action in Light of Gender Differences in Competitiveness. *Management Science 59*(1), 1–16.

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

Niederle, M. and L. Vesterlund (2011). Gender and competition. *Annual Review in Economics 3*, 601–630.

A Appendix Tables

Table A1: Descriptive Statistics

	mean	sd	min	max
A: Men				
Female Partner	0.52	0.51	0	1
Number of Boxes Collected in Risk Taking	38.74	12.46	10	66
Self-Reported Risk Taking	5.59	2.66	0	9
Self-Reported Competitiveness	6.80	2.36	1	10
N	44			
B: Women				
Female Partner	0.55	0.50	0	1
Number of Boxes Collected in Risk Taking	40.64	11.64	11	61
Self-Reported Risk Taking	4.45**	2.35	1	9
Self-Reported Competitiveness	5.36***	2.68	0	10
N	47			

T test for the Panel A and B mean comparison: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2: Performance, self-rank, and competition choices

	Men	Women
Task-1 performance	8.32 (0.58)	8.15 (0.45)
Task-2 performance	10.43 0.60)	9.74 (0.47)
Task-3 performance	10.91 (0.66)	10.51 (0.53)
Task-4 performance	11.27 (0.61)	10.47 (0.47)
Mean performance	10.23 (0.58)	9.72 (0.43)
Self-rank in Task-2	1.34 (0.07)	1.47 (0.04)
Competition chosen in Task-3	70% (0.07)	30% (0.07)***
Competition preferred in Task-4	70% (0.07)	28% (0.07)***
WTC in Task-4	2.11 (0.58)	-1.39 (0.50)***
N	44	47

Standard errors are reported in parentheses. T test for the mean comparison for men versus women: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A3: Performance, self-rank, and competition choices: gender composition

	Male Partner		Female Partner	
	Men	Women	Men	Women
Task-1 performance	8.10 (0.95)	8.52 (0.56)	8.52 (0.71)	7.85 (0.69)
Task-2 performance	10.62 (1.07)	9.57 (0.64)	10.26 (0.62)	9.88 (0.69)
Task-3 performance	11.10 (1.13)	10.24 (0.59)	10.74 (0.75)	10.73 (0.85)
Task-4 performance	11.24 (1.07)	10.52 (0.67)	11.30 (0.66)	0.42 (0.66)
Mean performance	10.26 (1.02)	9.71 (0.53)	10.21 (0.62)	9.72 (0.65)
Self-rank in Task-2	1.52 (0.11)	1.43 (0.11)	1.17 (0.08)	1.50 (0.10)***
Competition chosen in Task-3	67% (0.10)	29% (0.10)***	74% (0.09)	31% (0.09)***
Competition preferred in Task-4	67% (0.11)	29% (0.10)***	74% (0.09)	27% (0.09)***
WTC in Task-4	1.48 (0.81)	-1.74 (0.67)***	2.70 (0.82)	-1.12 (0.73)***
N	21	21	23	26
	Men		Women	
	Male Partner	Female Partner	Male Partner	Female Partner
Task-1 performance	8.10 (0.95)	8.52 (0.71)	8.52 (0.56)	7.85 (0.69)
Task-2 performance	10.62 (1.07)	10.26 (0.62)	9.57 (0.64)	9.88 (0.69)
Task-3 performance	11.10 (1.13)	10.74 (0.75)	10.24 (0.59)	10.73 (0.85)
Task-4 performance	11.24 (1.07)	11.30 (0.66)	10.52 (0.67)	0.42 (0.66)
Mean performance	10.26 (1.02)	10.21 (0.62)	9.71 (0.53)	9.72 (0.65)
Self-rank in Task-2	1.52 (0.11)	1.17 (0.08)***	1.43 (0.11)	1.50 (0.10)
Competition chosen in Task-3	67% (0.10)	74% (0.09)	29% (0.10)	31% (0.09)
Competition preferred in Task-4	67% (0.11)	74% (0.09)	29% (0.10)	27% (0.09)
WTC in Task-4	1.48 (0.81)	2.70 (0.82)	-1.74 (0.67)	-1.12 (0.73)
N	21	23	21	26

Standard errors are reported in parentheses. T test for the mean comparison for men versus women and those with a male partner versus a female partner: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A4: Choosing to compete in task-3: probit

	(1)	(2)	(3)
Female subject	-1.023***	-0.816	-0.781
	(0.34)	(0.51)	(0.50)
Female partner		0.376	0.449
		(0.53)	(0.49)
Female subject × female partner		-0.387	-0.472
		(0.74)	(0.70)
Task-2 performance	0.139*	0.145*	0.156**
	(0.07)	(0.07)	(0.07)
Improved Performance	0.052	0.054	0.052
(between task-1 and task-2)	(0.08)	(0.08)	(0.08)
Task-2 self-rank	0.259	0.146	
(first=1, second=0)	(0.38)	(0.41)	
Risk preference	0.360***	0.372***	0.379***
(self reported)	(0.08)	(0.08)	(0.08)
Constant	-2.923***	-3.174***	-3.259***
	(0.84)	(0.92)	(0.89)
N	91	91	91
chi2	54.133	54.644	54.519

Standard errors reported in parentheses. Significant levels are: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A5: Willingness to compete in task-4: OLS

	(1)	(2)	(3)
Female subject	-2.427***	-1.937**	-1.891**
	(0.64)	(0.96)	(0.94)
Female partner		1.535	1.617*
		(0.94)	(0.89)
Female subject × female partner		-0.958	-1.059
		(1.31)	(1.25)
Task-2 performance	0.218**	0.239**	0.251***
	(0.10)	(0.10)	(0.09)
Improved Performance	-0.050	-0.041	-0.041
(between task-1 and task-2)	(0.14)	(0.14)	(0.14)
Task-2 self-rank	0.561	0.210	
(first=1, second=0)	(0.72)	(0.76)	
Risk preference	0.773***	0.802***	0.812***
(self reported)	(0.13)	(0.13)	(0.12)
Constant	-4.751***	-5.712***	-5.799***
	(1.23)	(1.36)	(1.32)
N	91	91	91
r2	0.490	0.509	0.509

Standard errors reported in parentheses. Significant levels are: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

B Instructions

Screen 1

Good morning.

Thank you for participating to this experiment.

You are now taking part in an experiment in which you can earn between €5 and €30. It is therefore important that you read the instructions carefully as they explain how your decisions will affect your earnings.

The main part of the experiment will consist of four tasks. For each task you will receive specific instructions. These instructions will explain how you make decisions and how your decisions and the decisions of other participants influence your earnings. Therefore, it is important that you read the instructions carefully.

At the end of the experiment you will receive €5 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. At the end of the experiment we determine which task counts for payment by drawing a number between 1 and 4.

You will also participate to a complementary task that will be paid too.

Your earnings will be expressed in €. They will be paid to you in private at the end of the experiment.

Interactions are strictly anonymous. No one other than the experimenter will be informed about your decisions in the experiment.

From now on you are not allowed to communicate in any other way than specified in the instructions. Stay focused: do not use cellular phones or personal computers; otherwise, you may be excluded from the experiment.

Please carefully read these instructions.

Should you have any questions, please raise your hand and call the administrator.

Screen 2

Your personal characteristics

You are: a man / a women

Your age: xx

Your education level : Bac+1, +2, +3, +4, +5, +6 or more

Screen 3

Task 1 – Piece Rate

For Task 1 you will be asked to calculate the sum of five randomly chosen two-digit numbers. Example: 28+17+81+78+74=?.

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?

Screen 4

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. 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.

However for this task your payment depends on your performance relative to that of another participant, selected at random among the participants present in this room. You will be informed about his/her personal characteristics (gender, age, education level), and he/she will be informed about your personal characteristics, before the beginning of the task.

If Task 2 is the one randomly selected for payment, then your earnings depend on the number of problems you solve compared to your partner. The individual who correctly solves the largest number of problems will receive €1.00 per correct problem, while the other participant receives no payment. We refer to this as the **tournament payment**.

You will not be informed of how you did in the tournament until the end of the session.

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?

Screen 5

Personal characteristics of your partner

From now on, you will be paired with this person. His/her personal characteristics are:

A man / a women

Age: xx

Education level : xx

Screen 6

Task 3 – Choice of payment system

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. 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.

This time you can chose the compensation you want to have applied to this task: either "piece rate" or "tournament".

In the piece rate, you will receive €0.50 per problem solved correctly.

In the tournament, your performance will be evaluated relative to the performance of your partner in the Task 2 -tournament. The Task 2-tournament is the one you just completed. If you correctly solve more problems than your partner did in Task 2, then you receive two times the payment from the piece rate, which is €1.00 per correct problem. You will receive no earnings for this task under the tournament scheme if do not solve more problems correctly now, than your partner did in the Task-2 tournament.

You will not be informed of how you did in the tournament until the end of the session.

If there are ties, the winner will be randomly determined.

ARE THERE ANY QUESTIONS BEFORE WE BEGIN?

Screen 7

Task 4 - Choice of payment system (II)

As in the previous three tasks, you will be given 5 minutes to calculate the correct sum of a series of five 2-digit numbers. 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.

As in the task 3, the compensation scheme can be piece rate or tournament. In the piece rate, you will receive €0.50 per problem solved correctly.

In the tournament, your performance will be evaluated relative to the performance of your partner in the Task 2 -tournament. The Task 2-tournament is the one you just completed. If you correctly solve more problems than your partner did in Task 2, then you receive two times the payment from the piece rate, which is €1.00 per correct

problem. You will receive no earnings for this task under the tournament scheme if do not solve more problems correctly now, than your partner did in the Task-2 tournament.

We now invite you to think which of the two compensation systems is your preferred one. Suppose you can execute the task 4 and be compensated according to your preferred system. Is there a sum of money against you accept to renounce to this payment scheme and be compensated according to the other one ?

You will be allowed to indicate us your choices by means of a dynamic table with two columns and multiple rows. On a row you can chose between Option A – your payoff is the piece rate compensation plus a fixed amount of money a, and Option B - your payoff is the tournament compensation plus a fixed amount of money b. If this option is selected, this fixed amount (a or b) will be paid to you regardless of your performance in the execution of the task. On each row, the text in green indicates that the option is preferred to the option with the text in black.

The amounts (a, b) vary between €0.00 and €6.00 by increments of €0.50.

By placing the slider around the middle row in the table (the original position of the slider), you are indicating that you don't have a very strong preference between the Piece-rate or the Tournament compensation in executing the Task. By moving the slider up, you are indicating a stronger preference for the Tournament (you require a higher compensation if you had to play the Piece rate) and by moving the slider down, you are indicating a stronger preference for the Piece rate compensation (you require a higher compensation if you had to play the tournament).

Once that you expressed your preferences by positioning the slider, the computer will draw at random one of the rows in the table. You will be assigned to your preferred option (text in green) on that row in executing Task 4.

If at the end of the experiment Task 4 will be selected for payment, you will be compensated according to your performance in the task plus the fixed amount specified in that row.

As said, you will be asked to answer to such questions regarding your own choice, by using a dynamic table with two columns and multiple rows.

Screens 8-9

Two examples of how the slider works, about a hypothetical individual. In case of wrong answer subjects can retry once, then the Administrator will come and explain.

ARE THERE ANY QUESTIONS BEFORE WE BEGIN?

Screen 10

The subject indicates his preference on a Table (see main text for the table) by moving the slider up/down, than validates his/her choice.

The computer draws a row, and the subject is informed about the system of remuneration that will be applied to him/her in task 4.

Screen 11

The first part of the experiment is over now. Before moving to part two, please make a guess about you rank in the task-2 tournament. You will be paid €1.00 for a correct answer.

Your rank in Task 2 was: first / second

Screen 12

The box collecting task

On the PC screen you will see a field composed of 100 numbered boxes.

You earn 20 cents for every box that is collected. Every second, a box is collected, starting from the top left corner. Once collected, the box disappears from the screen, and your earnings are updated accordingly. At any moment you can see the amount earned up to that point.

Such earnings are only potential, however, because behind one of these boxes a time bomb is hidden that destroys everything that has been collected. You do not know where the time bomb is. You only know that it can be in any place with equal probability. Moreover, even if you collect the bomb, you will not know it until the end of the experiment. Your task is to choose when to stop the collecting process. You do so by hitting 'Stop' at any time.

At the end of the experiment, we will randomly determine the number of the box containing the time bomb by means of a bag containing 100 numbered tokens.

If you happen to have collected the box in which the time bomb is located, you will earn zero. If the time bomb is located in a box that you did not collect, you will earn the amount of money accumulated when hitting 'Stop'.

We will start with a practice round. After that, the paying experiment starts.

Screen 13.

Survey

1. How do you see yourself? Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?

0 [unwilling to take risks] ——— 10 [fully prepared to take risks]

2. How do you see yourself? Do you like to compete, or your rather prefer to avoid competition?

0 [I do not like to compete at all] ——— 10 [I really enjoy to compete]

Screen 14.

Do you remember the gender of your partner? yes/no

if yes: -> His/her gender was: male/female

Screen 14.

Results and payoff

The experiment is over now.

Your performance in Task 1:

Your performance in Task 2: number of correct answers [] and rank []

Your performance in Task 3: number of correct answers [] and compensation scheme []

Your performance in Task 4: number of correct answers [] and compensation scheme []. The fixed amount [].

The task selected for payment was [].

Your earnings are for this task are: []

The guess of your rank in task 2 brings you []

The box collecting task brings you []

Total earnings for the session: []

Thank you for your participation!

PARIS — SINGAPORE

ESSEC Business School
3 avenue Bernard-Hirsch
CS 50105 Cergy
95021 Cergy-Pontoise Cedex
France
Tel. +33 (0)1 34 43 30 00
www.essec.edu

ESSEC Executive Education
CNIT BP 230
92053 Paris-La Défense
France
Tel. +33 (0)1 46 92 49 00
www.executive-education.essec.edu

ESSEC Asia-Pacific
5 Nepal Park
Singapore 139408
Tel. +65 6884 9780
www.essec.edu/asia

ESSEC Africa-Atlantic
Plage des Nations
Sidi Bouknadel
Rabat-Salé
Morocco
Tel. +212 (0)5 30 10 40 19
www.essec.edu

ESSEC Africa-Indian Ocean
Royal Road, Pierrefonds
Mauritius
Tel. +230 401 2400
www.essec.edu
www.icsia.mu

MOROCCO — MAURITIUS

Contacts

Centre de Recherche
+33 (0)1 34 43 30 91
research.center@essec.fr

AACSB ACCREDITED
EFMD EQUIS ACCREDITED
affilié à la
CCI PARIS ILE-DE-FRANCE

ESSEC BUSINESS SCHOOL