DISCUSSION PAPER SERIES

IZA

IZA DP No. 9854

The Adverse Consequences of Tournaments: Evidence from a Field Experiment

Maria De Paola
Francesca Gioia
Vincenzo Scoppa

March 2016

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

The Adverse Consequences of Tournaments: Evidence from a Field Experiment

Maria De Paola
University of Calabria and IZA

Francesca Gioia
University of Edinburgh

Vincenzo Scoppa
University of Calabria and IZA

Discussion Paper No. 9854
March 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 9854
March 2016

ABSTRACT

The Adverse Consequences of Tournaments: Evidence from a Field Experiment*

We run a field experiment to investigate whether competing in rank-order tournaments with different prize spreads affects individual performance. Our experiment involved students from an Italian University who took an intermediate exam in which one part was awarded on the basis of their relative performance. Students were matched in pairs on the basis of their high school grades and each pair was randomly assigned to one of three different tournaments. Random assignment neutralizes selection effects and allows us to investigate if larger prize spreads increase individual effort. We do not find any positive effect of larger prizes on students' performance and in several specifications we do find a negative effect. Furthermore, we show that the effect of prize spreads on students' performance depends on their degree of risk-aversion: competing in tournaments with large spreads negatively affects the performance of risk-averse students, while it does not produce any effect on students who are more prone to take risks.

JEL Classification: J33, J31, J24, D81, D82, C93

Keywords: rank-order tournaments, incentives, prize spread, risk-aversion, randomized experiment

Corresponding author:

Maria De Paola
Department of Economics, Statistics and Finance
University of Calabria
Via Ponte Bucci
87036 Arcavacata di Rende (CS)
Italy
E-mail: m.depaola@unical.it

* We are grateful to the Dean of the Faculty of Economics at the University of Calabria for allowing us to run the experiment. We would like to thank Sabrina Giordano, Silvia Marchesi, Michela Ponzo and seminar participants at the University of Calabria for useful comments and suggestions. Francesca Gioia gratefully acknowledges financial support from Economic and Social Research Council.

1. Introduction

In many employment relationships, pay-for-performance schemes are used to motivate and select individuals. One of the problems faced by firms in implementing performance related pay contracts is associated to the availability of adequate measures of performance and to the costs of these measures.

Rank-order tournaments – which reward agents based on their relative performance rather than on their absolute performance (Lazear and Rosen, 1981) – have the advantage of relying on easily available measures since it is sufficient for the principal to rank individuals' performance rather than to measure precisely each single outcome. In addition, tournaments allow to filter out disturbance shocks that are common to all contestants and when these shocks tend to prevail tournaments may result more efficient than individual independent contracts (Green and Stokey, 1983).[1] However, tournaments are also affected by a number of problems: risks of sabotage of rivals' performance; inadequate incentives when agents have heterogeneous abilities; collusion among agents to exert low effort. These disadvantages are probably less relevant than the advantages since most employment relationships are characterized by competition among employees for bonuses or career advancements.

One of the main predictions of tournament theory is that larger spreads between high and low rewards are associated with better agents' performance due to both an incentive effect (contestants provide more effort) and a selection effect (high ability individuals self-select in these kinds of tournaments).

Disentangling these two effects empirically using observational data is not possible since individuals characterized by higher unobservable productivity tend to sort in high stake tournaments. As a consequence, researches have relied on both laboratory and field experiments that allow for a random assignment of individuals to tournaments with different prizes. While the evidence from laboratory experiments tends to confirm the main predictions of tournament theory, the evidence from the field shows that effort is not always sensitive to the structure of prizes. As argued by Leuven *et al.* (2011), the difference between lab and field results might be due to the fact that individuals involved in laboratory experiments are engaged in short-term tasks and, having to spend a certain amount of time in the laboratory, do not have the opportunity to devote their time to alternative tasks. Instead, individuals participating in a field experiment are engaged in longer-term activities and have the opportunity to choose between several alternative uses for their time.

An aspect that has received little attention in the literature on tournaments is the relationship between risk attitudes and prize spreads. When agents are risk-averse, larger prize spreads could be counterproductive because they entail higher risks for agents. Thus, in some contexts it could be optimal to limit the prize spread at the cost of lowering the levels of effort (Nalebuff and Stigliz, 1983).[2] A further

[1] Two further advantages of tournaments consist in: a) representing a commitment device to enforce principal's payments in contexts in which the principal can behave opportunistically because performance is not verifiable; b) forcing managers to make decisions on agents' rewards, since they cannot pay everyone the same.

[2] Millner and Pratt (1991) and Cornes and Hartley (2012) show that the relationship between risk aversion and agents' effort can be either positive or negative and depends on the third derivative of the utility function.

source of uncertainty arises from the nature of tournaments because they require agents to think strategically about their co-workers' decisions. As explained by Bull, Schotter and Weigelt (1987): "the agents will have to take into account two sources of uncertainty. One of these is a part of the existing theoretical literature, namely, the distribution of the prizes induced by the randomness in production [..]. The second source of uncertainty, which is not in the literature, is precisely the uncertainty concerning how the specific tournament that the agent enters will be played". In line with this prediction, lab experiments tend to find that more risk-averse individuals tend to avoid tournaments (Eriksson, Teyssier and Villeval, 2009; Dohmen and Falk, 2011).

In this paper we rely on a field experiment to shed more light on the incentive effects of larger prizes in tournaments and on whether and how these effects are related to agents' risk attitudes. Our experiment has involved a sample of Italian undergraduate students who undertook an intermediate test composed by two parts: one evaluated on the basis of student's absolute performance (piece rate) and the other evaluated on the basis of the student's relative performance with respect to a randomly chosen colleague (tournament). Students were matched in pairs on the basis of a measure of ability and each pair was randomly assigned to one of three types of tournaments with different prize structure: in the first tournament the prize spread between the winner and the loser was 6 points, in the second tournament the spread was 4 points and in the third tournament the spread was only 2 points. The random assignment of participants to different types of tournament allows us to investigate incentive effects of prize spreads isolating the related selection effects. Since our aim was to investigate the long-term incentive effects of tournaments, the assigned treatment and the parts of the program evaluated on the basis of absolute and relative performance, respectively, were communicated to students during the first week of teaching classes.

With respect to lab experiments, the advantages of our experiment are that individuals are studied in their natural environment when they perform real-effort tasks and we can observe individuals' reactions when they are able to choose between several alternative uses of their time. In addition, tournaments should be the ideal incentive scheme in our setting where the uncertainty deriving from common shocks (such as teaching material, instructors' quality, difficulty of test questions) is probably more important than the uncertainty deriving from agents' idiosyncratic shocks (shocks during the preparation of the exam, feeling better or worse the day of the exam, etc).

In our empirical analysis we do not find any positive effect of larger prizes on students' performance in the tournament and in some specifications we do find a negative effect. The effect of prize spreads on students' performance depends on their degree of risk-aversion: competing in tournaments with high spreads negatively affects the performance of risk-averse students while it does not produce any effect on students who are more prone to take risks.

The paper is organized as follows. Section 2 reviews the existing literature. In Section 3 we describe the experiment, present the data and conduct some balance checks. In Section 4 we carry out our main empirical analysis. In section 5 we study the relationship between the performance of students in the tournaments and their degree of risk-aversion. Section 6 concludes.

2. Related Literature

Our paper is related to the empirical literature testing tournament predictions. This literature relies both on observational data (mainly from sports) and on lab experiments.

Observational data typically do not allow to disentangle incentive and sorting effects. For instance, in a well-known study, Eriksson (1999) finds that the wage gap increases when one moves up in the hierarchy, but this is consistent with both incentives and sorting.[3] The difficulty to separate these two channels is common also to works relying on sports data. This strand of the literature starts with the seminal paper by Ehrenberg and Bognanno (1990) who, using data from professional golf tournaments, analyze the relationship between players' score in a tournament and the total available monetary prizes. Controlling for players' ability and opponents' quality, the difficulty of the course and weather conditions, they find that the level of prizes affects players' performance in line with theoretical predictions. Similar results are found for auto racing (Becker and Huselid, 1992), tennis (Sunde, 2009) and bowling (Abrevaya, 2002). However, the measures of abilities available do not allow to properly handle selection issues (which might be related to unobservable features) and then to understand whether the positive relationship between prize size and performance is due to an increase in effort or to the selection of players characterized by high unobservable ability in high stakes tournaments (or to both).

Instead, experimental studies, randomly assigning individuals to tournaments, avoid self-selection problems and are able to disentangle the incentive effects. Results found from laboratory experiments tend to confirm theory's predictions. The work by Bull, Schotter, and Weigelt (1987) is one of the first laboratory experiment testing tournament theory. Their results (replicated in many other lab experiments)[4] show that the average effort levels in tournaments are well predicted by theory (and the levels of effort provided by agents in tournaments and in piece rate schemes are similar). However, a very high variance in effort levels emerges in tournaments.

The evidence on tournaments found in field experiments is rather mixed. No incentive effects are found by Leuven *et al.* (2011) in a field experiment aimed at disentangling incentives and selection effects in tournaments. In their experiment students from a University in the Netherlands attending a course in microeconomics are firstly asked to choose between three tournaments offering different prizes: low (€1,000), medium (€3,000) and high (€5,000). Once students have chosen their favorite type of tournament and three groups corresponding to the three different tournaments are formed, students in each group are randomly assigned to a treatment group that actually competes for the prize and to a control group that does not take part in the competition. The comparison of the performance of the students assigned to the three control groups shows a strong selection effect: students with the best

[3] Studies based on firm data typically do not directly test the incentive and sorting effects of tournaments but offer some indirect tests based on the structure of compensation over the hierarchical ladder.

[4] For a review see Charness and Kuhn (2011) and Dechenaux, Kovenock and Sheremeta (2014). A large number of lab experiments on rank-order tournaments has tested a number of theoretical aspects. Schotter and Weigelt (1992) find that in tournaments among unequal agents handicaps can improve performance; Harbring and Irlenbusch (2008) show that large prize spreads can raise sabotage; Gill and Prowse (2012) find a discouragement effect for agents that are behind.

performance are those who chose to participate in the tournaments that offered the highest prize. To estimate the incentive effect, the authors compare for each selected tournament the performance of treated and control students and find no evidence of incentive effects. One possible explanation for the absence of incentive effects is that the expected probability of winning was very small (typically, more than 50 students competed for each prize).

In an experiment conducted by De Paola, Scoppa and Nisticò (2012) students enrolled at an Italian public university are randomly assigned to a control group and to two treatment groups in which students compete for the attainment of a small (€250) or a large (€700) monetary prize on the basis of their academic performance. The authors find that treated students increase academic performance but, surprisingly, large and small rewards produce very similar effects; furthermore, the positive effects are limited to high-ability students.

Bandiera, Barankay and Rasul (2005) compare piece rate and relative performance pay for fruit-pickers at a large firm. They find that workers' productivity is 50% higher in piece rate schemes, probably because socially connected workers in tournaments tend to provide low effort in order not to impose negative externalities on others or as a result of collusive behavior among agents. Fershtman and Gneezy (2011) in a field experiment involving schoolchildren in a short distance foot-race find that participants tend to perform better in high reward tournaments but the probability of quitting during the race is also high in these kinds of tournaments.[5]

The effect of risk-aversion on tournament outcomes has been investigated exclusively relying on lab experiments. Millner and Pratt (1991) find that more risk-averse individuals exert lower levels of effort than less risk-averse subjects. This result has been replicated in a number of other experiments (Anderson and Freeborn, 2010; Sheremeta and Zhang, 2010; Sheremeta, 2011; Shupp et al., 2013). Different approaches, ranging from observational data to lab and field experiments, have instead been used to investigate the relationship between risk attitudes and tournament participation (see for instance Eriksson, Teyssier and Villeval, 2009; Dohmen and Falk, 2011; Buser, Niederle and Oosterbeek, 2014). Overall these studies show that more risk-averse individuals tend to avoid tournaments.

3. The Experimental Design and the Data

3.1. Design and procedure

We run a field experiment involving 378 students enrolled in the academic year 2014-2015 at the courses of Microeconomics, Macroeconomics and Econometrics, offered by the First and Second Level Degree Course in Business and Administration at the University of Calabria.[6] These courses were taught to

[5] Some papers has investigated tournaments among teams. For instance, Hong, Hossain and List (2014) conduct a field experiment at a large Chinese manufacturing company offering a weekly prize to the team of workers with the highest per-hour productivity. They find a strong incentive effect of tournaments. Similar results are found by Delfgaauw et al. (2013) who analyze a sales competition between grocery stores belonging to the same retail chain.

[6] The University of Calabria is a middle-sized public university located in the South of Italy. It has currently about 32,000 students enrolled in different Degree Courses and at different levels of the Italian University system. Since

students during the second semester (teaching period from February to June). Each course is worth 10 credits, corresponding to 60 hours of teaching and to nominal 250 hours of study. For each course, all students attended the lectures in the same room, at the same time and with the same instructor and teaching material.

At the beginning of the courses, students were invited to join the experiment, in alternative to take the exam in the standard way at the end of the course.[7] In the experiment students take an intermediate exam divided into two parts. The first part included questions or exercises covering about 75% of the teaching material, and students' evaluation was based on their absolute performance ("piece rate"), for a maximum of 25 points. The second part ("tournament") was composed of 10 questions/exercises, covering the remaining 25% of the teaching material, and students were evaluated on the basis of their relative performance with respect to a randomly chosen colleague. Since the experiment concerns a real exam that matters for students' academic career we were not free to structure the whole test as a tournament.

Students joining the experiment were matched in pairs, within each course, on the basis of their high school grades and each pair was randomly assigned to one of three tournament typologies: in Tournament A or "High Spread Tournament", the student with the highest mark obtained 8 points while the student scoring the lowest mark obtained 2 points (with a prize spread of 6 points); in Tournament B or "Medium Spread Tournament " the student with the highest mark obtained 7 points while the student with the lowest mark obtained 3 points (with a prize spread of 4 points); in Tournament C or "Low Spread Tournament" the student with the highest mark obtained 6 points while the student with the lowest mark obtained 4 points (with a prize spread of 2 points). The points obtained in the tournament were added to the points gained in the piece rate part of the test.

In order to mimic the case of a firm changing the structure of prizes but paying the same total wage to its employees and to guarantee equal opportunity to all students, the three tournaments were structured in order to have the same expected prize (5 points).

Points awarded in the tournament were conditional on the reaching of a minimum threshold of 20% of correct answers. To avoid the risk of collusion among students, we did not allow for draws.[8] In case one student in a pair did not show up at the tournament, the competitor was randomly paired with another student with the same expected ability and the same type of tournament.

In order to complete the exam students joining the experiment had to take a second test at the end of the course covering the remaining teaching material. The exam grade was given by the average marks

---

the 2001 reform, the Italian University system is organized around three main levels: First Level Degrees (3 years of legal duration), Second Level Degrees (2 years more) and Ph.D. Degrees. In order to gain a First Level Degree students have to acquire a total of 180 credits. Students who have acquired a First Level Degree can undertake a Second Level Degree (acquiring 120 more credits). After having accomplished their Second Level Degree, students can enroll in a Ph.D. degree.

[7] The standard exam was structured as in the previous years and consisted of a single exam composed by questions and exercises and covering the whole course program to be taken at the end of the course. It was evaluated on the basis of student's absolute performance worth a maximum of 30 points with a minimum passing line of 18.

[8] In case of students with the same score within a pair, a second more careful evaluation was made by the instructor in order to establish a single winner.

obtained at the two tests. The final test was conducted following the standard rules without any competition among students and we do not analyze the related outcomes.

As required by the university administration for ethical reasons, students were free to join the experiment and, after joining it, were free to leave it at any point (after having registered for the experiment or after having taken the intermediate test) and to sit the standard exam. All the rules of the experiment were explained to students. They were given one week of time to choose whether to join to the experiment or to sit the standard exam. To join the experiment students had to fill out an on-line form in which they were also asked to answer a short survey on their family background, on their risk preferences and on the expected grade. Students were reassured that their answers would not have been considered before tests were graded.

A total of 378 students (about 78% of the 484 students enrolled in the 3 courses) decided to join the experiment.[9] Once obtained the list of participating students, within each course, we matched them in pairs on the basis of their high school grade (divided in 8 groups). The aim was to put in competition students with similar abilities, since tournaments with heterogeneous agents are predicted to provide weak incentives both for high abilities agents (quite confident to win) and for "underdogs" (that would be discouraged). Then, we stratified the pairs by course attended and high school grade (divided into quartiles) and randomly assigned them to the three different tournaments. The procedure assigned 124 students to Tournament A, 128 to Tournament B and 126 to Tournament C.

The pairs of students and the assigned tournament structure were immediately communicated to students in the classrooms and on the courses' webpages. Together with this information we also told students the chapters of the course program belonging to the piece rate and to the tournament part, respectively. This because we are interested in investigating the long term component of incentive effects, that is, how different prizes affect the allocation of students' effort to the study of the material of each part in the period going from the announcement of the tournament to the day in which the test was held. We are not interested in the immediate incentive effect of tournaments (i.e. if students exert more effort on the day of the test according to the prizes of the tournament).

The test was held on 18th April 2015 and so students had almost two months before their exam to allocate their effort taking into account the proposed incentive structure. 77 students did not show up at the intermediate test (20% of 378 students initially joining). Thus, 301 students took the exam: 102 in Tournament A, 103 in Tournament B and 96 in Tournament C.

The test questions regarding both parts of the intermediate exam were distributed simultaneously at the beginning of the test and students had 105 minutes to complete the whole test. Questions were the same for all the students in each course. Assistant professors marking the tests had no information on the treatment status of students.

[9] Estimation results from a Linear Probability Model (not reported) on the full sample of students enrolled in the courses show that younger students and students with a higher High School Grade are significantly more likely to join the experiment. Gender does not affect the decision to take the experiment in a statistically significant way.

3.2. Descriptive Statistics and Balance Checks

The design of the experiment produced three subsamples of students: those who enrolled in the courses (484), those who joined the experiment (378) and those who actually showed up at the intermediate test (301).

In Table 1 we provide descriptive statistics separately for the three subsamples of students. About 49% of students attending the courses are women, 49% represents also the proportion of female students joining the experiment while 52% of women took the test. Enrolled students are on average 22.2 years old, but they are somewhat younger in the experiment and among those taking the test (21.7).

Students enrolled in the courses obtained an average *High School Grade* of 82.9 (*High School Grade* ranges between 60 and 100); *High School Grade* becomes higher among students joining the experiment (83.2) and even more among students taking the test (85.1), suggesting a selection of better students in the test. About 50% of students in all three subsamples have studied in a Lyceum. About 53% of enrolled students are resident in the same area in which the University is located (*Same Area* is a dummy equal to one for students coming from the province of Cosenza and 0 otherwise), while this percentage decreases to, respectively, 52% and 50% for students joining the experiment and students actually showing up at the test.[10]

Students' answers to the on-line survey filled-in when joining the experiment allow us to build a self-reported measure of risk attitude. The question we asked is the same as in the German Socio-Economic Panel (see Dohmen *et al.*, 2011) and is formulated as follows: "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? Please tick a box on the scale, where the value 1 means: `not at all willing to take risks' and the value 10 means: `very willing to take risks'." We build the variable *Risk Propensity*[11] simply reflecting the values chosen by students. *Risk Propensity* is on average 6.5 for the subsample of students joining the experiment and it is almost the same for students showing up at the test.[12]

In the survey, we also asked students about their parents' education. *Parents' education* represents the average years of education of students' father and mother. In the sample of students joining the experiment the mean is 11.7. It is slightly lower (11.6) for students taking the exam.

Treatment groups were evenly balanced in the subsample of students joining the experiment (about 33% in each type of tournament) and remain balanced also in the subsample of students taking the

[10] In the final sample, about 43% of students are students of Microeconomics, 39% are students of Macroeconomics and 17% study Econometrics.

[11] To elicit students' attitudes towards risk we also asked students a second question regarding the preferred amount to invest in a risky asset. Specifically we asked: "We would like to ask you a hypothetical question that you should answer as if the situation were a real one. You are offered the opportunity of invest an amount permitting you, with the same probability, either to double the investment or to lose all the capital invested. Which amount you choose to invest? [100,000; 80,000; 60,000; 40,000; 20,000; 0]. Then, we built the variable *Risky Investment* equal to 6 for an investment of 100,000, 5 for an investment of 80,000 and so on. This variable is highly correlated to *Risk Propensity* (ρ=0.60) In the following econometric analyses we use only *Risk Propensity* but we obtain very similar results using the second measure based on the amount to invest (estimates not reported but available upon requests).

[12] Consistently with Dohmen *et al.* (2011) we find that the willingness to take risks is negatively correlated to student's age and high school grade. Women turn out to be significantly more risk-averse. We do not find instead any relationship between parents' education and risk-aversion.

intermediate test (about 34% in the *High* and *Medium Spread Tournament* and 32% in the *Low Spread Tournament*).

Table 1. Descriptive Statistics. Mean and SD

	Enrolled at the courses	Joining the Experiment	Taking the Exam
Female	0.488	0.487	0.522
	(0.500)	(0.500)	(0.500)
Age	22.236	21.757	21.788
	(2.920)	(2.601)	(2.640)
High School Grade	82.934	83.249	85.110
	(10. 964)	(11.088)	(10.812)
Lyceum	0.492	0.503	0.492
	(0.500)	(0.501)	(0.501)
Same Area	0.535	0.524	0.502
	(0.499)	(0.500)	(0.501)
Micro	0.417	0.471	0.432
	(0.494)	(0.500)	(0.496)
Macro	0.423	0.370	0.395
	(0.423)	(0.484)	(0.490)
Econometrics	0.159	0.159	0.173
	(0.159)	(0.366)	(0.379)
Risk Propensity		6.513	6.432
		(1.808)	(1.840)
Parents' Education		11.712	11.571
		(3.268)	(3.124)
High Spread Tournament		0.328	0.339
		(0.470)	(0.474)
Medium Spread Tournament		0.339	0.342
		(0.474)	(0.475)
Low Spread Tournament		0.333	0.319
		(0.472)	(0.467)
Tournament Performance			4.122
			(3.283)
Standard Grade			12.481
			(7.367)
Skipped Questions			3.659
			(3.647)
Mistakes			2.230
			(2.066)
Observations	484	378	301

Notes: Standard Deviations are reported in parentheses.

Tournament Performance represents the score obtained by students in the competition (ranging from 0 to 10) and is our main dependent variable. Students obtained on average 4.1; the SD is quite high (3.3) and almost 20% of students obtain a score of zero. *Standard Grade* is the score students obtained in the piece rate part of the exam and it ranges from 0 to 25. On average students score 12.5 points out of 25; about 25% of students obtain a score of 7 or less.

For the tournament section we have also calculated how many points students miss because they leave blank the answers (*Skipped Questions*) or because they make mistakes (*Mistakes*). On average students lost 3.7 point because they did not attempt to answer the questions and 2.2 points due to incorrect answers.

To investigate the effects produced by prize spreads in tournaments we need comparable individuals in the three treatment groups. However, the possibility of students to switch to the standard exam after knowing the assigned treatment could have invalidated our random assignment to treatments by introducing a self-selection element in the sample of students who effectively took the exam. Thus, we verify if groups are balanced both among students joining the experiment and among students showing up at the test.

In columns (1) to (3) of Table 2 we report means for a number of individual characteristics for the three treatment groups in the subsample of students joining the experiment (Panel A) and in the subsample of students showing up at the test (Panel B). Columns (4) to (6) report the differences of means and *t*-stats of tests of equality of variables' means for the three pairs of treatments. In the last column, we report the *F*-stats and *p*-values for a test of equality of variables' means across all three groups.

In both subsamples, we are not able to reject the hypothesis that the randomization was successful in creating comparable treatment groups as regards the observable characteristics: there are no significant differences between the treatment status in terms of students' gender, *Age*, *High School Grade*, type of High School attended, *Same Area*, *Parents' Education* and *Risk Propensity*.

Table 2. Students' Characteristics across Treatment Groups

Panel A. Students joining the experiment							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	High Spread T.	Medium Spread T.	Low Spread T.	(1) vs. (2) (*t*-stat)	(1) vs. (3) (*t*-stat)	(2) vs. (3) (*t*-stat)	F-stat (*p*-value)
Female	0.508	0.453	0.500	0.054	0.08	-0.047	0.44
				(0.871)	(0.127)	(-0.746)	(0.642)
Age	21.653	21.720	21.898	-0.067	-0.245	-0.177	0.30
				(-0.211)	(-0.742)	(-0.530)	(0.744)
High School Grade	83.387	83.133	83.230	0.254	0.157	-0.097	0.02
				(0.184)	(0.110)	(0.070)	(0.983)
Lyceum	0.548	0.492	0.468	0.056	0.080	0.024	0.84
				(0.891)	(1.26)	(0.380)	(0.432)
Same Area	0.532	0.500	0.540	0.032	-0.007	-0.039	0.23
				(0.510)	(-0.117)	(-0.631)	(0.798)
Parents' Education	12.105	11.539	11.500	0.566	0.605	0.039	1.34
				(1.43)	(1.410)	(0.095)	(0.262)
Risk Propensity	6.403	6.445	6.690	-0.042	-0.287	-0.245	0.92
				(-0.188)	(-1.25)	(-1.06)	(0.397)
Obs.	124	128	126				

Panel B. Students taking the exam							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	High Spread T.	Medium Spread T.	Low Spread T.	(1) vs. (2) (*t*-stat)	(1) vs. (3) (*t*-stat)	(2) vs. (3) (*t*-stat)	F-stat (*p*-value)
Female	0.549	0.505	0.510	0.044	0.039	-0.005	0.23
				(0.631)	(0.542)	(-0.078)	(0.792)
Age	21.675	21.821	21.873	-0.146	-0.197	-0.051	0.15
				(-0.400)	(-0.534)	(-0.133)	(0.861)
High School Grade	84.618	85.408	85.312	-0790	-0.695	-0.095	0.16
				(-0.525)	(-0.438)	(0.063)	(0.852)
Lyceum	0.539	0.485	0.448	0.054	0.091	0.037	0.83
				(0.767)	(1.283)	(0.528)	(0.436)
Same Area	0.520	0.485	0.500	0.034	0.019	-0.014	0.12
				(0.487)	(0.274)	(-0.204)	(0.887)
Parents' Education	12.029	11.553	11.104	0.476	0.925	0.449	2.19
				(1.10)	(2.029)	(1.039)	(0.114)
Risk Propensity	6.314	6.378	6.614	-0.065	-0.301	-0.236	0.73
				(-0.260)	(-1.127)	(-0.892)	(0.485)
Obs.	102	103	96				

Note. In columns (1), (2) and (3) are reported variable means. In columns (4), (5) and (6) we report differences of means and the corresponding t-stat in parentheses. In column (7), we report the *F*-statistic and *p*-value for a test of equality of variable means across all three groups.

As a further check, in Appendix A (Table A1) we also analyze the probability of taking the exam on the sample of students joining the experiment. We find that the decision to take the test is not affected by the assigned type of tournament and is not related to students' risk propensity.

4. Students' Performance in Relation to Prize Spreads

In this section we investigate the incentive effects of tournaments with different prize spreads. By virtue of random assignment to different types of tournaments we know that participants have the same observable and unobservable characteristics and thus we are able to exclude that selection effects are driving our results.

We analyze the effect of prize spreads by estimating several specifications of the following OLS model:

$$Tournament\ Performance_i = \beta_0 + \beta_1\left(High\ Spread\ Tournament_i\right) + \beta_2\left(Medium\ Spread\ Tournament_i\right) + \beta_3 X_i + \varepsilon_i$$

where the dependent variable, $Tournament\ Performance_i$, is the total score that the student obtains in the competition part of the test; $High\ Spread\ Tournament_i$ is a dummy variable for students assigned to the tournament with the highest spread (8 points for the winner and 2 for the loser); $Medium\ Spread\ Tournament_i$ is a dummy variable for students in the tournament with the medium spread (7 points for the winner and 3 for the loser); students assigned to the tournament with the lowest spread (6 points for the winner and 4 for the loser) are left as the reference category: β_1 and β_2 represent the causal effect in terms of student's performance of being assigned to the high spread and medium spread tournament with respect to the low spread tournament, respectively; X_i denotes the vector of student's predetermined characteristics and cognitive abilities; ε_i is an error term.

Table 3 reports OLS estimates for the impact of price spreads on students' academic performance. In all our regressions Standard Errors are corrected for heteroskedasticity. The first specification controls only for the impact of the two tournaments with the largest and medium spreads (and for dummies of course attended). It emerges that neither being assigned to the tournament with the highest spread nor to the tournament with the medium spread produce any effect on students' performance: the coefficients are negative but far from being statistically significant. The second specification adds among controls students' gender and age. Women tend to perform better while older students achieve worse grades. The impact of prize spreads is still statistically not significant.

In the third specification we add some measures of students' abilities: *High School Grade*, *Lyceum* and *Parents' Education* and control for students' area of residence. While the measures of abilities tend to affect positively the score obtained at the competition (in particular, 10 points more of *High School Grade* increase by 1.37 the performance at the competition, an effect that corresponds to about 0.4 Standard Deviations of the dependent variable). Again, the effects of competing in tournaments with larger prize spreads are negative but not statistically significant.[13]

In the specification in column (4) we also control for students' self-reported risk attitude (*Risk Propensity*). Adding this control does not change our main results: taking part in tournaments paying larger prizes does not affect students' performance. The attitude towards risk of students does not seem to affect directly their performance.

Since individuals in tournaments take strategic decisions that could be affected by the characteristics of the competitors, in column (5) we control for some opponents' characteristics: gender,

[13] We also run a specification controlling for student's *Expected Grade* (a measure of self-reported ability). The direct effect is positive and statistically significant but the impact of type of tournament does not change.

Age and *Lyceum*.[14] None of them turns out to be statistically significant and in addition the impact of tournament's type is still not significant.

Table 3. Prize Spreads and Students' Performance. OLS Estimates

	(1)	(2)	(3)	(4)	(5)
High Spread T.	-0.263	-0.387	-0.342	-0.355	-0.353
	(0.458)	(0.451)	(0.415)	(0.419)	(0.417)
Medium Spread T.	-0.089	-0.131	-0.162	-0.171	-0.180
	(0.459)	(0.448)	(0.405)	(0.407)	(0.411)
Macro	-0.211	0.030	-0.284	-0.266	-0.204
	(0.414)	(0.418)	(0.390)	(0.388)	(0.410)
Econometrics	1.380***	3.294***	1.804**	1.810**	2.233**
	(0.529)	(0.750)	(0.841)	(0.839)	(1.107)
Female		1.207***	0.528	0.519	0.544
		(0.373)	(0.349)	(0.349)	(0.356)
Age		-0.373***	-0.310***	-0.316***	-0.320***
		(0.097)	(0.107)	(0.107)	(0.106)
High School Grade			0.137***	0.136***	0.135***
			(0.016)	(0.017)	(0.017)
Lyceum			0.462	0.474	0.488
			(0.331)	(0.331)	(0.336)
Parents' Education			0.052	0.051	0.057
			(0.058)	(0.058)	(0.059)
Same Area			0.176	0.170	0.139
			(0.326)	(0.328)	(0.331)
Risk Propensity				-0.038	-0.046
				(0.095)	(0.095)
Female Rival					-0.096
					(0.341)
Rival's Age					-0.057
					(0.111)
Rival's Lyceum					-0.388
					(0.343)
Constant	4.087***	11.225***	-2.025	-1.576	0.042
	(0.418)	(2.063)	(2.632)	(2.870)	(3.841)
Observations	301	301	301	301	301
Adjusted R-squared	0.018	0.078	0.250	0.248	0.244

Notes: The dependent variable is *Tournament Performance*. Standard errors (corrected for heteroskedasticity) are reported in parentheses. The symbols ***, **, * indicate that the coefficients are statistically significant at the 1, 5 and 10 percent level, respectively.

By examining students' achievement on the exam, it is easy to realize that a large number of students took the exam without effectively providing an adequate level of effort. For example, 57 students (19%) obtained a score of 0 in the competition part, while 40% of students obtain 10 point or less (out of 25) in the piece-rate part. In the Italian system it is usual that students take the exam more than once (students are allowed to take the exam up to 5 times in an academic year) and so their incentives to study are not so strong. The effects of our treatments could be diluted by the presence of this kind of students, since they tend to study little regardless of the assigned treatment.

[14] Given that pairs of students were formed on the basis of their high school grades, the correlation between own and *Rival's High School Grade* is 0.98. For this reason we do not control for this rival's characteristic. If we include *Rival's High School Grade* results are very similar but the two *High School Grade* variables are not significant because of their high collinearity.

Therefore, we focus on students that we assume studied for the exam, not just came to attempt it, by excluding students that obtain a *Standard Grade* below 13 (these students would not have passed the exam adding an expected score of 5 in the competition part) and run the same regressions of Table 3 on this sample of 144 students.[15] Estimates are reported in Table 4. Now we find that in general students assigned to the *High Spread Tournament* tend to perform worse than students assigned to the *Low Spread Tournament* – from 0.7 to 0.9 points according to the specification. The corresponding *p*-values are around 0.05-0.13.

Furthermore, the performance of students in the *Medium Spread Tournament* is about 0.6 points lower than the reference category, although this effect is rather imprecisely estimated (*p*-values are around 0.15). Overall, the evidence suggests that larger spreads tend to worsen students' performance.

We do not find statistically significant differences between men and women according to prize spreads, although, with respect to men, women seem to perform worse in *High Spread Tournaments* but better in *Medium Spread Tournaments*.[16]

[15] Whereas in the whole sample 57 students (19%) obtain a score of 0 in the competition part, in this restricted sample only 1 student scores 0. Changing slightly the threshold of 13 has little effect on our estimates.
[16] For similar findings in a related experiment see De Paola, Gioia and Scoppa (2015).

Table 4. Prize Spreads and Students' Performance. Excluding Students with Low Scores. OLS Estimates

	(1)	(2)	(3)	(4)	(5)
High Spread T.	-0.665	-0.922*	-0.932*	-0.903*	-0.761
	(0.492)	(0.472)	(0.483)	(0.486)	(0.497)
Medium Spread T.	-0.522	-0.692	-0.659	-0.660	-0.656
	(0.490)	(0.471)	(0.463)	(0.466)	(0.462)
Macro	-1.468***	-0.954**	-1.059**	-1.129**	-1.073**
	(0.461)	(0.455)	(0.442)	(0.436)	(0.471)
Econometrics	-0.683	2.116***	1.539**	1.562**	1.404
	(0.605)	(0.720)	(0.755)	(0.755)	(1.149)
Female		0.175	0.004	0.006	-0.015
		(0.403)	(0.415)	(0.412)	(0.405)
Age		-0.605***	-0.538***	-0.534***	-0.556***
		(0.104)	(0.115)	(0.115)	(0.118)
High School Grade			0.052**	0.054**	0.064***
			(0.023)	(0.023)	(0.023)
Lyceum			0.293	0.239	0.242
			(0.392)	(0.392)	(0.399)
Parents' Education			0.041	0.034	0.029
			(0.074)	(0.075)	(0.082)
Same Area			0.292	0.294	0.405
			(0.384)	(0.385)	(0.390)
Risk Propensity				0.125	0.130
				(0.120)	(0.119)
Female Rival					-0.749*
					(0.390)
Rival's Age					0.047
					(0.137)
Rival's Lyceum					0.110
					(0.436)
Constant	7.668***	20.069***	13.381***	12.474***	-0.761
	(0.446)	(2.267)	(3.656)	(3.820)	(0.497)
Observations	144	144	144	144	144
Adjusted R-squared	0.050	0.149	0.163	0.163	0.167

Notes: The dependent variable is *Tournament Performance*. Standard errors (corrected for heteroskedasticity) are reported in parentheses. The symbols ***, **, * indicate that the coefficients are statistically significant at the 1, 5 and 10 percent level, respectively.

Alternatively, instead of focusing on a specific subsample of students, we use the whole sample and control for the grade students have obtained in the piece rate part of the exam as a further measure of abilities. We run our main regressions and report our estimates in Table 5. Ten additional points of *Standard Grade* increase performance in the competition part by 3.5 points (*t*-stat=19). Standard errors are almost halved thanks to this very good measure of student's abilities. More importantly, we find that *High Spread Tournament* causes a decrease in the performance of about 0.5 points, significant at the 5 or 10 percent level. In these specifications, we find that *Medium Spread Tournament* does not produce any effect on students' performance.

Table 5. Controlling for Standard Grade. OLS Estimates

	(1)	(2)	(3)	(4)	(5)
High Spread T.	-0.546*	-0.584**	-0.573**	-0.534*	-0.482*
	(0.280)	(0.281)	(0.276)	(0.279)	(0.279)
Medium Spread T.	0.026	0.008	-0.015	0.015	0.028
	(0.271)	(0.269)	(0.263)	(0.265)	(0.263)
Standard Grade	0.350***	0.341***	0.316***	0.320***	0.321***
	(0.014)	(0.015)	(0.016)	(0.016)	(0.016)
Macro	-0.301	-0.214	-0.279	-0.336	-0.337
	(0.242)	(0.263)	(0.260)	(0.258)	(0.273)
Econometrics	0.610	1.309**	0.991	0.963	0.807
	(0.376)	(0.637)	(0.667)	(0.673)	(0.851)
Female		0.440*	0.290	0.314	0.318
		(0.240)	(0.244)	(0.242)	(0.246)
Age		-0.133*	-0.135	-0.117	-0.113
		(0.079)	(0.082)	(0.084)	(0.084)
High School Grade			0.047***	0.048***	0.051***
			(0.013)	(0.013)	(0.013)
Lyceum			0.110	0.070	0.085
			(0.220)	(0.220)	(0.222)
Parents' Education			0.035	0.036	0.035
			(0.039)	(0.039)	(0.040)
Same Area			0.327	0.347	0.373*
			(0.224)	(0.224)	(0.225)
Risk Propensity				0.117*	0.120*
				(0.069)	(0.069)
Female Rival					-0.299
					(0.226)
Rival's Age					0.031
					(0.074)
Rival's Lyceum					-0.124
					(0.231)
Constant	-0.060	2.583	-1.470	-2.841	-3.614
	(0.262)	(1.673)	(2.047)	(2.173)	(2.763)
Observations	301	301	301	301	301
Adjusted R-squared	0.633	0.639	0.655	0.658	0.657

Notes: The dependent variable is *Tournament Performance*. Standard errors (corrected for heteroskedasticity) are reported in parentheses. The symbols ***, **, * indicate that the coefficients are statistically significant at the 1, 5 and 10 percent level, respectively.

Since *Tournament Performance* is a limited variable ranging between 0 to 10, we also experiment using a Tobit estimator, allowing for left-censored observations at a score of 0 and for right-censored observations at a score of 10. We run the same specifications of Table 5. The negative effect of *High Spread Tournament* is confirmed in Tobit estimates (results not reported and available upon request).

As a further check we also use the prize spread in linear form, building a variable *Spread* equal to 6 for individuals in *High Spread Tournament*, 4 for those in *Medium Spread Tournament* and 2 for those in *Low Spread Tournament*. When we control for *Spread* instead of the two treatment dummies we find that *Spread* has a negative and significant impact (coefficient is around -0.14, with a t-stat of -2.1) on students' tournament performance (estimates not reported).

In principle, the assignment of students to the three different tournaments might have affected their performance not only at the competition part of the test, but also at the piece rate section. Students could have reacted to the prize structure by substituting effort from one section of the test to the other. To analyze this aspect we replicate specifications reported in Table 3 considering as a dependent variable

Standard Grade. In all specifications (reported in Appendix B, Table B1), we do not find any statistically significant treatment effect on students' performance in the piece rate section, suggesting that students have not been affected by prize spreads in deciding how much to study for the piece rate part of the test.

5. Risk-Aversion and Incentive Effects of Tournaments

In this section we investigate whether the responses of students to tournaments with different prize spreads depend on their attitudes toward risk. The direct effect of larger prize spreads is that they increase the variance of earnings and therefore reduce the expected utility of risk-averse agents. A second effect, the so-called "strategic uncertainty" – pointed out by Bull *et al.* (1987) – is that since tournaments require that individuals behave strategically and take into account how competitors will play, the risks are magnified if agents are, realistically, uncertain about competitors' abilities, or their rationality, risk-attitude or cost-of-effort, etc., and are thus uncertain about how opponents will play the game. Risk-averse agents will suffer more from this further source of uncertainty.

In order to empirically analyze the relationship between risk-aversion and performance in tournaments, we use the measure of students' risk-aversion that we elicited in the preliminary survey and we include in our main regressions the variable *Risk Propensity* and two interaction terms between the latter and the two dummies *High Spread Tournament* and *Medium Spread Tournament*.

The corresponding OLS estimates are reported in Table 6. In column (1) we only control for dummies of the attended course. In column (2) we control also for *Female* and *Age*. In column (3) we control for *High School Grade*, *Lyceum*, *Parents' Education* and *Same Area* while in column (4) we include competitor's characteristics (Gender, Age, Lyceum). The coefficients of control variables are not reported to save space.

Table 6. Risk Aversion and Students' Performance in Tournaments. OLS Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
High Spread T.	-3.220**	-3.168**	-2.588*	-2.478*	-1.655**	-1.595**
	(1.541)	(1.500)	(1.393)	(1.400)	(0.752)	(0.756)
Medium Spread T.	-3.396**	-2.968*	-2.536*	-2.459*	-0.993	-0.990
	(1.570)	(1.560)	(1.437)	(1.445)	(0.800)	(0.807)
High Spread T.* Risk Propensity	0.448*	0.421*	0.343*	0.325		
	(0.231)	(0.227)	(0.200)	(0.200)		
Medium Spread T.* Risk Propensity	0.503**	0.430*	0.362*	0.348*		
	(0.233)	(0.231)	(0.206)	(0.206)		
Risk Propensity	-0.426***	-0.407***	-0.262*	-0.258*		
	(0.146)	(0.143)	(0.135)	(0.133)		
Risk Taker					-1.312*	-1.300*
					(0.705)	(0.702)
High Spread T.* Risk Taker					1.881**	1.794**
					(0.889)	(0.894)
Medium Spread T.* Risk Taker					1.149	1.136
					(0.932)	(0.935)
Observations	301	301	301	301	301	301
Adjusted R-squared	0.030	0.088	0.252	0.247	0.255	0.251

Notes: The dependent variable is *Tournament Performance*. Standard errors (corrected for heteroskedasticity) are reported in parentheses. Controls (not reported) as in Table 3. The symbols ***, **, * indicate that the coefficients are statistically significant at the 1, 5 and 10 percent level, respectively.

Interestingly, we find that the effect of prize spreads on students' performance depends on their propensity to undertake risks. Results are very similar in all the specifications. For students fully averse to risk (*Risk Propensity* equal to 0) we find a very strong negative effect both for *High Spread Tournament* and for *Medium Spread Tournament*: students' performance is about 3 points lower when assigned to these two types of tournaments (statistically significant at the 5 percent level). However, the interaction terms *High Spread Tournament*(Risk Propensity)* and *Medium Spread Tournament *(Risk Propensity)* are both positive (around 0.3-0.4) and statistically significant, implying that for an individual with an average risk propensity of 6 the impact of both treatment groups is negative but close to zero (not statistically significant). For example, in column (3) we show that the impact of being assigned to the *High Spread Tournament* is –0.53 (=–2.588+(6*0.343)) (t-stat=–1.22).

To check the robustness of our results to different functional forms of our measure of *Risk Propensity* in columns (5) and (6) of Table 6 we replicate specification (3) and (4) but we use a dummy variable *Risk Taker*, taking the value of one for values of *Risk Propensity* higher than 6 (and the value of zero otherwise), and we interact *Risk Taker* for our two treatment dummies. We find a negative and statistically significant effect of competing in *High Spread Tournament* for students who are risk-averse (–1.655), significant at the 5 percent level, while this effect becomes almost zero (0.226=–1.655+1.881) with a *t*-stat of 0.46 for individuals more prone to take risks. Being assigned to the *Medium Spread Tournament* produces a negative but not statistically significant effect (–0.993, *t*-stat=–1.24) for risk-averse individuals. The impact of *Medium Spread Tournament* shrinks again to almost zero (0.156=–0.993+1.149; *t*-stat=0.33) for more risk-prone students.

Furthermore, we use *Spread* in linear form and interact it with *Risk Propensity* (estimates not reported). We find that one point of spread decreases performance of about 0.7 (significant at the 5 percent level) for risk-averse individuals whereas has almost a zero effect for more risk prone individuals.

In Table 7 we report the same estimates of Table 6 but include *Standard Grade* among our control variables. The estimates are very similar, although the impact of *Medium Spread Tournament* tends to be smaller in magnitude and generally not significant.

Table 7. Risk Aversion and Students' Performance in Tournaments. Controlling for Standard Grade. OLS Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
High Spread T.	-2.785**	-2.796**	-2.586**	-2.508**	-1.315**	-1.255**
	(1.120)	(1.113)	(1.118)	(1.123)	(0.573)	(0.573)
Medium Spread T.	-1.493	-1.370	-1.223	-1.113	-0.074	-0.009
	(1.086)	(1.087)	(1.125)	(1.114)	(0.602)	(0.585)
High Spread T.* Risk Propensity	0.350**	0.347**	0.317**	0.312*		
	(0.162)	(0.161)	(0.159)	(0.160)		
Medium Spread T.* Risk Propensity	0.235	0.213	0.188	0.173		
	(0.156)	(0.156)	(0.159)	(0.159)		
Risk Propensity	-0.088	-0.085	-0.044	-0.033		
	(0.096)	(0.096)	(0.102)	(0.102)		
Risk Taker					-0.114	-0.056
					(0.508)	(0.505)
High Spread T.* Risk Taker					1.143*	1.146*
					(0.658)	(0.665)
Medium Spread T.* Risk Taker					0.086	0.018
					(0.672)	(0.660)
Observations	301	301	301	301	301	301
Adjusted R-squared	0.640	0.645	0.661	0.660	0.660	0.659

Notes: The dependent variable is *Tournament Performance*. Standard errors (corrected for heteroskedasticity) are reported in parentheses. Controls (not reported) as in Table 3. The symbols ***, **, * indicate that the coefficients are statistically significant at the 1, 5 and 10 percent level, respectively.

To take into account the censored nature of our dependent variable we estimate specifications in Table 7 with a Tobit estimator, allowing for censoring at score 0 and 10. The qualitative results are very similar to the previous OLS estimates (not reported).

All these findings suggest consistently that risk-averse individuals tend to perform worse in tournaments with large spreads.[17] Since in our setting individuals have to take jointly the piece rate and tournament sections of our exam, risk-averse students who dislike the high uncertainty in rewards embedded in large spread tournaments have probably chosen to substantially give up from competition, saving on the provision of effort in the tournament stage, but they still come to the test because of the presence of the piece-rate section that accounts for 75% of the exam.[18] Probably these risk-averse agents would have simply chosen to not participate in the tournament if this choice was possible.

To better investigate this aspect, we use *Skipped Questions* (number of points lost due to students not attempting the questions) and *Mistakes* (points lost due to writing incorrect answers) for the tournament section. If individuals choose to not participate in competition we should observe that the skipped questions at the tournament part of the test are more frequent for risk-averse individuals in large spreads' tournaments. In Table 8 we report the same specifications of Table 7 using *Skipped Questions* as

[17] In line with our results, Ariely *et al.* (2009) have found that workers faced with high monetary rewards in individual performance-contingent-pay tend to decrease their productivity.

[18] In fact, we do not find that more risk-averse students tend to leave the experiment if they are assigned to tournaments with larger spreads.

the dependent variable. As expected, we find that in high and medium spread tournaments students with high risk-aversion tend to leave blank the answers (2-3 more points missed than the *Low Spread Tournament*), while more risk-prone individuals miss fewer points. Although the statistical significance of many coefficients is not high, the qualitative findings are pretty clear.

Table 8. Skipped Questions (Missed Points). OLS Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
High Spread T.	2.451	2.492	2.001	2.137	1.614**	1.679**
	(1.759)	(1.749)	(1.583)	(1.556)	(0.764)	(0.784)
Medium Spread T.	3.113*	2.665	2.401	2.620	1.164	1.314
	(1.769)	(1.801)	(1.695)	(1.686)	(0.805)	(0.830)
High Spread T.* Risk Propensity	-0.319	-0.313	-0.245	-0.243		
	(0.273)	(0.271)	(0.238)	(0.230)		
Medium Spread T.* Risk Propensity	-0.406	-0.334	-0.290	-0.314		
	(0.272)	(0.274)	(0.253)	(0.249)		
Risk Propensity	0.390**	0.353**	0.225	0.246*		
	(0.158)	(0.157)	(0.148)	(0.144)		
Risk Taker					1.426**	1.524**
					(0.689)	(0.705)
High Spread T.* Risk Taker					-1.739*	-1.607*
					(0.965)	(0.963)
Medium Spread T.* Risk Taker					-0.903	-1.040
					(1.011)	(1.020)
Observations	301	301	301	301	301	301
Adjusted R-squared	0.004	0.037	0.181	0.188	0.188	0.194

Notes: The dependent variable is *Skipped Questions (Missed Points).* Standard errors (corrected for heteroskedasticity) are reported in parentheses. Controls (not reported) as in Table 3. The symbols ***, **, * indicate that the coefficients are statistically significant at the 1, 5 and 10 percent level, respectively.

On the other hand, in Table 9 we use as a dependent variable *Mistakes*. In this case – as expected – all coefficients are pretty close to zero and we do not find any significant effect neither for the type of tournament nor for the degree of risk aversion.

Table 9. Mistakes (Missed Points). OLS Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
High Spread T.	0.763	0.670	0.583	0.336	0.043	-0.081
	(1.250)	(1.225)	(1.220)	(1.189)	(0.539)	(0.537)
Medium Spread T.	0.281	0.299	0.130	-0.164	-0.161	-0.312
	(1.099)	(1.077)	(1.083)	(1.066)	(0.589)	(0.589)
High Spread T.* Risk Propensity	-0.129	-0.108	-0.098	-0.082		
	(0.184)	(0.182)	(0.181)	(0.175)		
Medium Spread..* Risk Propensity	-0.096	-0.095	-0.071	-0.033		
	(0.160)	(0.158)	(0.160)	(0.156)		
Risk Propensity	0.035	0.052	0.036	0.011		
	(0.120)	(0.114)	(0.115)	(0.113)		
Risk Taker					-0.109	-0.218
					(0.472)	(0.471)
High Spread T.* Risk Taker					-0.151	-0.196
					(0.641)	(0.635)
Medium Spread T.* Risk Taker					-0.255	-0.105
					(0.686)	(0.680)
Observations	301	301	301	301	301	301
Adjusted R-squared	-0.002	0.011	0.006	0.035	0.008	0.039

Notes: The dependent variable is *Mistakes (Missed Points).* Standard errors (corrected for heteroskedasticity) are reported in parentheses. Controls (not reported) as in Table 3. The symbols ***, **, * indicate that the coefficients are statistically significant at the 1, 5 and 10 percent level, respectively.

6. Concluding Remarks

Using a field experiment involving about 300 students competing in pairs to obtain the best performance in rank-order tournaments with different prize spreads we have found that – in contrast with theoretical predictions – agents do not provide high effort when prize spreads are larger. When we take into account agents' risk aversion, we find that while the effect of large spreads is not significant for risk-prone individuals, it is strongly detrimental for the performance of risk-averse individuals.

The finding that risk-averse individuals perform worse in high spread tournaments is quite new in the literature. We were able to isolate incentive effects from selection effects thanks to the random assignment of agents to different types of tournaments. The experiment was also designed to avoid both collusion and heterogeneity in abilities among competitors since these factors could have reduced effort.

Leuven *et al.* (2011) showed that tournaments have no incentive effects but in their setting the probability of winning the prizes was quite low (about 2%): we strengthen their conclusions showing that not only incentive effects are null in a setting in which the probability of winning is 50%, but that they could even backfire for risk-averse individuals, inducing them to provide low effort as a defensive strategy when the uncertainty in rewards is predominant.

Our results are consistent with those emerging from some lab experiments: the problem of uncertainty in tournaments – arising mainly from the difficulties to think strategically and predict how competitors will play – was firstly pointed out by Bull, Schotter, and Weigelt (1987) – while Eriksson, Teyssier and Villeval (2009) and Dohmen and Falk (2011) show how risk-averse individuals tend to shy away from tournament schemes. In our framework, individuals have to take jointly the piece rate and tournament sections of the exam, but risk-averse students who dislike the risks embedded in large spread tournaments have probably chosen to provide less effort in the tournament stage.

Although in real labor markets workers typically tend to self-select in incentive schemes according to their characteristics, our strategy gives account of the incentive effects of tournaments in contexts in which imperfect mobility (due for example to search and matching frictions or slackness of local labor markets) prevents workers from sorting in the preferred incentive scheme or in settings where it takes time to move after the introduction of a new incentive scheme.

These findings suggest that the widespread use of tournaments in the form of promotions – rather than being inspired by the need to provide incentives – is perhaps more related to the selection effects and to the aim of sorting workers to jobs on the basis of their talents.

References

Abrevaya, J. (2002). Ladder tournaments and underdogs: Lessons from professional bowling. *Journal of Economic Behavior and Organization* 47, no. 1:87–101.

Anderson, L. R., and Freeborn, B. A. (2010). Varying the intensity of competition in a multiple prize rent seeking experiment. *Public Choice*, 143(1-2), 237-254.

Ariely, D., Gneezy, U., Loewenstein, G., and Mazar, N. (2009). Large stakes and big mistakes. The *Review of Economic Studies*, 76(2), 451-469.

Bandiera, O., Barankay, I., and Rasul, I. (2005). Social preferences and the response to incentives: Evidence from personnel data. *The Quarterly Journal of Economics*, 917-962.

Becker, B. E., and Huselid, M. A. (1992). The incentive effects of tournament compensation systems. *Administrative Science Quarterly* 37, no. 2:336–50.

Bull, C., Schotter, A., and Weigelt, K. (1987). Tournaments and piece rates: An experimental study. *Journal of Political Economy* 95, no. 1:1–33.

Buser, T., Niederle, M., and Oosterbeek, H. (2014). Gender, competitiveness and career choices. *Quarterly Journal of Economics*, 129(3), 1409–1447

Charness, G., and Kuhn, P. (2011). Lab labor: What can labor economists learn from the lab?. *Handbook of labor economics*, 4, 229-330.

Cornes, R., and Hartley, R. (2012). Risk aversion in symmetric and asymmetric contests. *Economic Theory*, 51(2), 247-275.

De Paola, M., Gioia, F., and Scoppa, V. (2015). Are females scared of competing with males? Results from a field experiment. *Economics of Education Review*, 48, 117-128.

De Paola, M., Scoppa, V., and Nisticò, R. (2012). Monetary incentives and student achievement in a depressed labor market: Results from a randomized experiment. *Journal of Human Capital*, 6(1), 56-85.

Dechenaux, E., Kovenock, D., and Sheremeta, R. M. (2014). A survey of experimental research on contests, all-pay auctions and tournaments, *Experimental Economics*, 1-61.

Delfgaauw, J., Dur, R., Sol, J., and Verbeke, W. (2013). Tournament incentives in the field: Gender differences in the workplace. *Journal of Labor Economics*, 31(2), 305-326.

Dohmen, T. and Falk, A. (2011). Performance Pay and Multidimensional Sorting: Productivity, Preferences, and Gender. *American Economic Review*, 101, 556-90.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3), 522-550.

Ehrenberg, R. G., and Bognanno, M. L. (1990). Do tournaments have incentive effects? *Journal of Political Economy* 98, no. 6:1307–24.

Eriksson, T. 1999. Executive compensation and tournament theory: Empirical tests on Danish data. *Journal of Labor Economics* 17, no. 2:262–80.

Eriksson, T., Teyssier, S., and Villeval, M. C. (2009). Self-selection and the efficiency of tournaments. *Economic Inquiry*, 47(3), 530-548.

Fershtman, C. and Gneezy, U. (2011). The Trade-off between Performance and Quitting in High-Power Tournaments. *Journal of the European Economic Association*, 9, 318-336.

Gill, D., and Prowse, V. (2012). A structural analysis of disappointment aversion in a real effort competition. *The American economic review*, 469-503.

Green, J. and. Stokey, N., 1983, "A Comparison of Tournaments and Contracts", *Journal of Political Economy* 91:3, 349-364.

Harbring, C., and Irlenbusch, B. (2003). An experimental study on tournament design. *Labour Economics* 10, no. 4:443–64.

Harbring, C., and Irlenbusch, B. (2008). How many winners are good to have?: On tournaments with sabotage. *Journal of Economic Behavior & Organization*, 65(3), 682-702.

Hong, F., Hossain, T., & List, J. A. (2015). Framing manipulations in contests: a natural field experiment. Journal of Economic Behavior & Organization, 118, 372-382.

Lazear, E. P., and Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy* 89, no. 5: 841–64.

Leuven, E., Oosterbeek, H., Sonnemans, J., and Van Der Klaauw, B. (2011). Incentives versus sorting in tournaments: Evidence from a field experiment. *Journal of Labor Economics*, 29(3), 637-658.

List, J., Van Soest, D., Stoop, J., and Zhou, H. (2014). On the role of group size in tournaments: Theory and evidence from lab and field experiments (No. w20008). National Bureau of Economic Research.

Millner, E. L., and Pratt, M. D. (1991). Risk aversion and rent-seeking: An extension and some experimental evidence. *Public Choice*, 69(1), 81-92.

Nalebuff, B. and Stiglitz, J. (1983), "Prizes and Incentives: Towards a General Theory of Compensation and Competition", *The Bell Journal of Economics*, Vol. 14, pp. 21-43.

Niederle, M. and Vesterlund, L., (2011), “Gender and Competition”, *Annual Review in Economics*, 3, 601–30.

Schotter, A., and Weigelt. K. (1992). Asymmetric tournaments, equal opportunity laws, and affirmative action: Some experimental results. *Quarterly Journal of Economics* 107, no. 2:511–39.

Sheremeta, R.M. (2011). Contest design: An experimental investigation. *Economic Inquiry*, 49, 573-590

Sheremeta, R.M. and Zhang, J. (2010). Can groups solve the problem of over-bidding in contests? *Social Choice and Welfare*, 35, 175-197

Shupp, R., Sheremeta, R. M., Schmidt, D., and Walker, J. (2013). Resource allocation contests: Experimental evidence. *Journal of Economic Psychology*, 39, 257-267.

Sunde, U. (2009). Heterogeneity and performance in tournaments: A test for incentive effects using professional tennis data. *Applied Economics* 41, no. 25:3199–3208.

Appendix A. Probability to take the test

As a further check aimed at reassuring that randomization was successful in the subsample of students effectively taking part in the experiment and that there was no self-selection, we analyze students' decision to show up at the test on the sample of students joining to the experiment (378). Estimation results from a Linear Probability Model are reported in Table A1. In the first specification, we only control for courses' dummies and type of tournaments. In the second column we add some students' individual characteristics (gender and age) and cognitive abilities (measured by the typology of high school attended and the *High School Grade*). In the third column we add *Parents' Education*, province of residence and *Risk Propensity*. In column (4) we add some characteristics of student's competitor (gender, Age, Lyceum).

We find that the dummies for *High Spread Tournament* and for *Medium Spread Tournament* (leaving as reference category *Low Spread Tournament*) are not significant in any specification and then the decision to take the test was not affected by the assigned type of tournament.

Furthermore, to check whether students in the three tournaments differ in terms of observable characteristics we also include among regressors interaction terms between the dummy variables *High Spread Tournament* and *Medium Spread Tournament* and the other control variables (not reported). We find that none of the estimated coefficients is statistically significant, thus students assigned to different treatments have similar observable characteristics. In particular, students' decision to show up at the test is not influenced by their risk attitude and students assigned to different treatments do not differ in terms of self-reported risk attitude.

Table A1. Determinants of Students' Decision to Take the Exam. Linear Probability Model

	(1)	(2)	(3)	(4)
High Spread T.	0.060	0.055	0.055	0.054
	(0.051)	(0.049)	(0.049)	(0.050)
Medium Spread T.	0.042	0.042	0.036	0.037
	(0.051)	(0.048)	(0.048)	(0.049)
Macro	0.119***	0.097**	0.097**	0.101**
	(0.045)	(0.046)	(0.046)	(0.048)
Econometrics	0.136**	0.119	0.117	0.179
	(0.055)	(0.101)	(0.099)	(0.128)
Female		0.024	0.012	0.013
		(0.041)	(0.042)	(0.042)
Age		-0.017	-0.017	-0.018
		(0.014)	(0.014)	(0.014)
High School Grade		0.011***	0.011***	0.010***
		(0.002)	(0.002)	(0.002)
Lyceum		-0.021	-0.007	-0.009
		(0.039)	(0.041)	(0.041)
Parents' Education			-0.008	-0.008
			(0.007)	(0.007)
Same Area			-0.047	-0.052
			(0.039)	(0.039)
Risk Propensity			-0.012	-0.014
			(0.011)	(0.011)
Female Rival				0.044
				(0.041)
Rival's Age				-0.010
				(0.013)
Rival's Lyceum				-0.025
				(0.041)
Constant	0.697***	0.138	0.358	0.652
	(0.046)	(0.338)	(0.367)	(0.465)
Observations	378	378	378	378
Adjusted R-squared	0.017	0.111	0.115	0.113

Notes: The dependent variable is *Taking the Exam*. Standard errors (corrected for heteroskedasticity) are reported in parentheses. The symbols ***, **, * indicate that the coefficients are statistically significant at the 1, 5 and 10 percent level, respectively.

Appendix B. The Impact of Prize Spreads on Standard Grade

In Table B1 we report the same specifications as in Table 3 by considering as dependent variable *Standard Grade*. We do not find any effect on the Standard Grade of being assigned to different types of tournaments.

Table B1. The Impact of Prize Spreads on Standard Grade. OLS Estimates

	(1)	(2)	(3)	(4)	(5)
High Spread T.	0.808	0.575	0.732	0.559	0.402
	(1.054)	(1.044)	(0.964)	(0.949)	(0.949)
Medium Spread T.	-0.327	-0.408	-0.463	-0.582	-0.649
	(1.020)	(1.007)	(0.944)	(0.941)	(0.946)
Macro	0.257	0.716	-0.017	0.218	0.413
	(0.947)	(0.973)	(0.917)	(0.913)	(0.946)
Econometrics	2.198*	5.816***	2.574	2.646	4.445**
	(1.151)	(1.672)	(1.669)	(1.661)	(2.170)
Female		2.244***	0.754	0.643	0.705
		(0.857)	(0.815)	(0.814)	(0.814)
Age		-0.704***	-0.553***	-0.621***	-0.644***
		(0.209)	(0.207)	(0.205)	(0.202)
High School Grade			0.287***	0.275***	0.261***
			(0.037)	(0.037)	(0.037)
Lyceum			1.113	1.263	1.258
			(0.789)	(0.787)	(0.805)
Parents' Education			0.054	0.047	0.067
			(0.136)	(0.136)	(0.139)
Same Area			-0.479	-0.552	-0.732
			(0.767)	(0.762)	(0.765)
Risk Propensity				-0.485**	-0.515**
				(0.211)	(0.210)
Female Rival					0.634
					(0.812)
Rival's Age					-0.273
					(0.239)
Rival's Lyceum					-0.823
					(0.767)
Constant	11.838***	25.310***	-1.758	3.956	11.398
	(0.954)	(4.430)	(5.574)	(6.029)	(7.513)
Observations	301	301	301	301	301
Adjusted R-squared	0.002	0.042	0.190	0.201	0.201

Notes: The dependent variable is *Standard Grade*. Standard errors (corrected for heteroskedasticity) are reported in parentheses. The symbols ***, **, * indicate that the coefficients are statistically significant at the 1, 5 and 10 percent level, respectively.

Since students' performances at the two parts of the exam are closely related, as a robustness check we also allow for covariance between the error terms by jointly estimating the equations with, respectively, the *Tournament Performance* and the *Standard Grade* as dependent variables, using a generalized least squares estimator (Zellner's Seemingly Unrelated Regression, SURE), which is more efficient than estimating the two equations separately. Our results remain qualitatively unchanged (not reported): being assigned to tournaments with larger prize spreads does not affect student's performance neither at the tournament nor at the piece rate part of the test.