



The adverse consequences of tournaments: Evidence from a field experiment[☆]



Maria De Paola^{a,b}, Francesca Gioia^{c,*}, Vincenzo Scoppa^{a,b}

^a Department of Economics, Statistics and Finance, University of Calabria, Via Ponte Bucci, Arcavacata di Rende (CS), 87036, Italy

^b Institute for the Study of Labor (IZA), Bonn, Germany

^c School of Economics, University of Edinburgh, 30 Buccleuch Place, Edinburgh, EH8 9JT, United Kingdom

ARTICLE INFO

Article history:

Received 6 July 2017

Revised 5 April 2018

Accepted 4 May 2018

JEL classification:

J33

J31

J24

D81

D82

C93

Keywords:

Rank-order tournaments

Incentives

Prize spread

Risk-aversion

Randomized experiment

ABSTRACT

We ran 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 exam that was partly evaluated on the basis of 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 performance. 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 taking risks.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

Tournaments – which reward agents based on their relative performance (Lazear and Rosen, 1981) – are widely used in organizations: employees often compete with one another to get promoted or to receive some kind of reward. Among the main advantages of tournaments there is the fact that they rely on easily available albeit non verifiable measures of performance: it is sufficient for the principal to rank individuals' performance rather than measuring precisely each single

[☆] 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 the Editor of the Journal of Economic Behavior & Organization, Daniela Puzzello, three anonymous referees, Marco Bertoni, Giorgio Brunello, Davide Fiaschi, Tommaso Luzzati, Nicola Meccheri, Antonio Nicolò, Luca Nunziata, Michela Ponzo, Lorenzo Rocco, Mauro Sylos Labini, and seminar participants at the University of Calabria, Pisa and Padua and at the IWAEE Conference for useful comments and suggestions. Francesca Gioia gratefully acknowledges financial support from [Economic and Social Research Council](#).

* Corresponding author.

E-mail addresses: m.depaola@unical.it (M. De Paola), f.gioia@sms.ed.ac.uk (F. Gioia), v.scoppa@unical.it (V. Scoppa).

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 be more efficient than individual independent contracts (Green and Stokey, 1983).¹

The standard theory (Lazear and Rosen, 1981) suggests that in a tournament an increase in the prize spread raises the effort of competitors (typically assumed to be risk-neutral), hence their performance. However, there are also situations in which increasing the prize spread can be counterproductive. Falk et al. (2008) show that when outcomes are evaluated in relation to a reference point (Kahneman and Tversky, 1979), due to loss aversion, individuals might be less willing to work hard as the prize spread increases. High spreads can produce undesired effects also when agents are averse to risk. Larger spreads imply a raise in the variance of outcomes and this negatively affects the utility of risk averse subjects. If this is so, it could be optimal to limit the prize spread at the cost of lowering the levels of effort (Nalebuff and Stiglitz, 1983).

Empirical works testing the relationship between spread and effort have to deal with the fact that individuals characterized by higher unobservable productivity tend to sort into high stake tournaments. This implies that larger spreads are associated with the higher performance of such agents due to both an incentive effect (contestants provide more effort) and a selection effect (high ability individuals self-select in these kinds of tournaments). To disentangle these two effects researchers have relied on both laboratory and field experiments that, by using the random assignment of individuals to tournaments with different prizes, avoid sorting and allow researchers to identify the incentive effect. 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 can choose between alternative ways of spending their time.

In light of the scant and ambiguous results found in the literature, our paper aims to provide additional evidence on the incentive effects of tournaments that takes also into consideration agents' risk aversion. Despite the relevance of risk, the empirical literature investigating how the incentive effects of tournaments vary in relation to individual risk aversion is scant. The existing literature, relying both on lab and field experiments, has mainly focused on tournament participation and individual risk aversion, showing that risk averse individuals tend to avoid competitive settings. However, in many labor markets, mobility is imperfect and individuals might not be able to react to changes in the wage structure by changing their jobs. So, it is interesting to understand how risk attitudes affect individual reaction to tournaments involving higher or lower risks.

We answer this research question by running a field experiment with a sample of Italian undergraduates who undertook an intermediate test composed of two parts: one evaluated on the basis of the student's absolute performance (piece rate) and the other evaluated on the student's performance with respect to a randomly chosen colleague (tournament). Students were matched in pairs on the basis of ability and each pair was randomly assigned to one of three types of tournament with different prize structure: in the first tournament the prize spread between the winner and the loser was high, in the second tournament the spread was medium and in the third tournament the prize spread was low.

The random assignment of participants to different types of tournament allows us to disentangle a purely incentive effect of higher prize spreads. Also, since in all the three types of tournament competition is between pairs of similar ability, there is no reason for individuals assigned to different tournaments to change effort in relation to the expected ability of the pool of competitors. This implies that, in our setting, the incentive effects only derive from the prize size (and not from the fact that in high stake tournaments the expected ability of the opponent is higher). In addition, we focus on long term incentive effects in a context in which individuals can allocate their time to different activities. In fact, both the assigned treatment and the parts of the program evaluated on the basis of absolute and relative performance, respectively, have been communicated to students during the first week of classes, so that they had almost two months to make decisions about effort intensity and allocation. This is a clear advantage compared to lab experiments in which students' choice as regards the use of their time is limited. Finally, we collect information on individuals' risk preferences before they know the assigned treatment and participating in the tournament and therefore we are able to measure heterogeneity along this dimension without any reverse causation problem. While the relationship between risk aversion and prize spreads has been recently examined in a few lab experiments (Sheremeta and Wu, 2012), to the best of our knowledge, our experiment is the first in the field exploring this issue.

Results from our analysis show that students' performance does not increase when they compete in large prize tournaments if we average effects without considering individual risk-aversion. This evidence is in line with findings reported in Leuven et al. (2011) who also find no incentive effect of larger prize spreads in tournaments. However, we find that large prize spreads do effect some students' performance: 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. Although there is a good deal of attrition in our experiment, we are confident that our results capture important attitudes of risk-averse individuals towards high-spread tournaments.

¹ However, tournaments are also affected by a number of problems: risk 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 employees competing for bonuses or career advancement.

The paper is organized as follows. [Section 2](#) reviews the existing literature. In [Section 3](#) we describe the experiment, present the data and conduct 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](#) analyzes the performance in the piece-rate part of the exam. [Section 7](#) concludes.

2. Related literature

Our paper belongs to the empirical literature testing tournament predictions, which relies both on observational data (mainly from sports) and on lab experiments.

Observational data typically do not allow researchers 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.² The difficulty in separating these two channels is also common to work based 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' scores in a tournament and total prize money. 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 available measures of ability do not allow to properly handle selection issues (which might be related to unobservable features) and from there 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 but unobservable ability into high stakes tournaments (or to both).

In contrast to observational analyses, experimental studies, randomly assigning individuals to tournaments, avoid self-selection problems and are able to disentangle the incentive effects. Much of the literature relies on laboratory experiments using both chosen effort and real effort tasks. Results found from chosen effort experiments, where participants are given information on the costs and benefits of different effort options and are asked to state the level of effort they would like to exert, tend to confirm the theory's predictions. For instance, [Bull et al. \(1987\)](#), in one of the first laboratory experiments testing tournament theory, show that 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. Similar findings are shown in many other lab experiments on rank-ordered tournaments.³ Consistent with the theory, many experimental studies find that subjects increase their effort in response to an increase in the value of the prize ([Harbring and Irlenbusch, 2003](#); [Harbring and Lünser, 2008](#); [Falk et al., 2008](#); [Sheremeta and Wu, 2012](#)).

Mixed results instead are found from experiments using real-effort tasks, with some studies showing a positive relationship between effort and prize size and others showing a lack of response to incentives or even a negative effect ([Jenkins et al., 1998](#); [Gneezy and Rustichini, 2000](#); [Mohnen et al., 2008](#); [Ariely et al., 2009](#); [Araujo et al., 2016](#); [DellaVigna and Pope, 2017](#), forthcoming).

Mixed evidence also emerges from field experiments on tournaments.⁴ 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 first asked to choose between three tournaments offering different prizes: low (€1000), medium (€3000) and high (€5000). 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 performance are those who chose to participate in the tournaments offering 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 such effects is that the expected probability of winning was very low (typically, more than 50 students competed for each prize).

In an experiment conducted by [De Paola et al. \(2012\)](#) 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) 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.⁵

² 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 along the hierarchy.

³ For a review see [Charness and Kuhn \(2011\)](#) and [Dechenaux et al. \(2015\)](#). A large number of lab experiments on rank-order tournaments has tested various 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 the risk of sabotage; [Gill and Prowse \(2012\)](#) find a discouragement effect for agents that are behind.

⁴ The mixed results found by field experiment and real effort experiments might depend on the fact that when individuals perform a real effort task additional aspects, such as enjoyment, can play an important role in determining effort (see [Erkal et al., 2018](#)).

⁵ This result is consistent with the theoretical prediction that in a highly discriminatory contest, or all-pay auction, the disadvantaged players drop out (see [Baye et al., 1996](#)). See also [Cason et al. \(2010\)](#).

Bandiera et al. (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 exert little 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 school children in a short distance 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.⁶

Research on the effect of risk-aversion on tournament outcomes⁷ has relied exclusively on lab experiments. A number of papers find that more risk-averse individuals exert lower levels of effort than less risk-averse subjects (Millner and Pratt 1991; Anderson and Freeborn, 2010; Sheremeta and Zhang, 2010; Sheremeta, 2011), while others do not find any significant effect of risk aversion (Shupp et al., 2013; Kräkel and Niekne, 2015).

The effect produced by risk aversion on the relationship between prize spreads and individual effort has been recently investigated by Sheremeta and Wu (2012). Relying on a chosen-effort experiment they show that the effect produced by changes in the prize spread and in the loser's prize are related to individual risk aversion. More risk averse agents tend to increase their effort when the loser's prize increases and the prize spread is held constant.

Different approaches, ranging from observational data to lab and field experiments, have been used instead to investigate the relationship between risk attitudes and tournament participation (see for instance Eriksson et al., 2009; Dohmen and Falk, 2011; Buser et al., 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.⁸ These courses were held during the second semester (from February to June). Each course was compulsory and worth 10 credits, corresponding to 60 h of teaching and to a nominal 250 h 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. As these courses are offered at different years of the Degree program there are no students who attended more than one course.⁹

At the beginning of the courses, we informed students that they could choose between two different exam schemes: the “standard exam” and an “alternative exam” scheme introducing competition between students through a relative performance evaluation.¹⁰

The standard exam consisted of a single three-hour test composed of questions and exercises covering the whole course program and was scheduled 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 pass of 18.

The alternative exam scheme, related to the experiment, was composed of two tests, each covering half of the program, to be taken right after the first half of the course (intermediate test) and at the end of it (final test), respectively. The intermediate test – which we take into consideration in our experiment – was divided into two parts. The first part included 12 questions/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 6 questions/exercises (for a maximum of 10 points), 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 student of similar ability. The test questions regarding both parts of the intermediate exam were distributed simultaneously at the beginning and students had 105 min to complete the whole test. Questions were the same for all the students in each course. Examiners marking the tests had no information on the treatment status of students. 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.

⁶ Some papers have investigated tournaments with teams. For instance, Hong et al. (2015) 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 stores belonging to the same footwear retail chain.

⁷ Millner and Pratt (1991) and Cornes and Hartley (2012) show that the relationship between risk aversion and effort can be either positive or negative and depends on the third derivative of the utility function. Hillman and Katz (1984) and Skaperdas and Gan (1995) show that risk aversion implies lower efforts in contests.

⁸ 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 the 2001 reform, the Italian University system is organized into three main levels: First Level Degrees (3 years of legal duration), Second Level Degrees (2 further years) 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 apply to enroll for a Ph.D.

⁹ Microeconomics and Macroeconomics are offered, respectively, in the first and second year of the First Level Degree Program. Since these courses are attended by a large number of students, classes are organized by dividing students into two groups depending on their surnames. In our experiment, we have considered the Microeconomics course attended by students with surnames from A to L and the Macroeconomics course attended by students with surnames from M to Z. Econometrics is offered at the Second Level Degree Program.

¹⁰ The instructions given to the students are reported in Appendix A.

The final test – to be taken at the end of the course to complete the exam in the alternative scheme – consisted of questions and exercises covering all the remaining teaching material and allowing students to score a maximum of 30 points on the basis of their absolute performance (with a time of 105 min).¹¹ The final grade obtained by those who participated in the experiment was given by the average of the grades obtained in the intermediate and final tests. Students scoring less than 18 did not pass the exam and had to retake the standard exam scheme.¹²

Students joining the experiment were matched in pairs, within each course, on the basis of their high school grades.¹³ The aim was to let students with similar abilities to compete against each other, since tournaments with heterogeneous agents are predicted (under the assumptions typically used to model the cost of effort)¹⁴ to provide weak incentives both to high abilities agents (confident to win) and to “underdogs” (discouraged). Then, 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 to determine the grade of the intermediate 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, in addition, to guarantee equal opportunity to all students, the three tournaments were structured so that all would compete for the same expected prize (5 points).

The bonus awarded in the tournament was conditional on reaching a minimum threshold of 20% of correct answers. To avoid the risk of collusion, we did not allow for draws.¹⁵ In the case of one of a pair not showing up at the tournament, the remaining student was randomly paired with another student with the same expected ability and the same type of tournament in order to determine the tournament prize.

As required by the university, for ethical reasons, students were free to join the experiment and, free to leave it at any point (after registering for the experiment or having taking the intermediate test) and to sit the standard exam. In the Italian University system, students are allowed to retake an exam as many times as they wish. However, students typically cannot retake an intermediate exam and, if they fail it, they have to take the standard exam (without the possibility to split the class workload in two parts).

One might wonder whether the incentives in our experimental exam were high enough to induce students to provide a high level of effort and obtain a good grade. This, however, is not a major concern for two reasons. First of all, retaking an exam has a significant cost for students, since their academic career is slowed down to retake an exam instead of studying for new exams. So, the chance to re-sit the exam does not imply that students have less incentive to study. Secondly, in general students much prefer to split the class workload into two parts by taking an intermediate exam and focusing only on one part of the program at a time.¹⁶ In the specific environment created by our experiment they are aware that if they failed to pass the intermediate exam they would have to retake the standard one.

All the rules of the experiment were explained to students and published on the course webpages (see [Appendix A](#)). They were given one week to choose whether to join the experiment or to take the standard exam. To join the experiment students had to fill out an on-line form in which they were also required to complete a short survey on their family background, their risk preferences and the expected grade at the exam. Students were assured that their answers would not be considered before the tests were graded.

A total of 378 students (about 78% of the 484 students enrolled in the 3 courses) decided to join the experiment.¹⁷ Once the list of participants from each course had been obtained, they were paired on the basis of their High School Grades (divided in 8 groups).¹⁸ The pairs were then stratified by degree course and high school grade (divided into quartiles)

¹¹ The final test, similarly to the standard exam, was conducted following the standard rules without any competition among students. For this reason we do not analyze the related outcomes.

¹² The structure of the intermediate and final tests is equivalent to that of the standard exam.

¹³ The High School (or secondary school) in Italy ends with a national examination (so called “Esame di Maturità”). The score obtained in this examination and the performance in the main subjects of the last three years in the High School are weighted to determine the final High School Grade, which ranges between 60 and 100. It is a comprehensive measure of ability of students. High School Grade tends to be a very reliable predictor of exam performance, see, for example, [De Paola and Scoppa \(2015\)](#) who consider a large sample of students enrolled at the same university considered in the present paper, or [Di Pietro \(2013\)](#) considering a large representative sample of all Italian High School graduates.

¹⁴ Even if the “common wisdom” prevailing in the literature is that contests with symmetric contestants maximize total effort, [Drugov and Ryvkin \(2017\)](#) show that in a general tournament model, it is possible to come up with parameterizations that predict an increase in total effort for heterogeneous players.

¹⁵ In the case of students with the same score within a pair, a second more careful evaluation was made by the instructor in order to establish the winner. If their performance was still identical between competitors, a winner was randomly selected.

¹⁶ [De Paola and Scoppa \(2011\)](#) show that students benefit a lot from having the class program tested through an intermediate examination covering the first part of the course material and a final exam covering exclusively the second part instead of a single exam covering the whole program.

¹⁷ 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 participate in a statistically significant way.

¹⁸ In all the courses, there was an even number of students joining the experiment.

and randomly assigned 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. We decided to reveal the identity of the competitors in order to reflect the common situation faced by individuals when they compete in the workplace in relative performance evaluations (for example, for a bonus or promotion) in which competitors tend to know each other.¹⁹

Together with this information the students were told which parts of the course program belonged to the piece rate and to the tournament part, respectively. This is because of our interest 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, from the announcement of the tournament to the day of the test, rather than in the immediate incentive effect of tournaments (i.e. if students exert more effort on the day of the test depending on the prizes).

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

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. About 49% of students attending the courses are female and this also represents the percentage of women who joined the experiment. This percentage rises to 52% when considering students who actually took the test. The average age of students is 22.2, while the average age of those joining the experiment and of those who actually took the test is slightly younger, 21.8.

Students enrolled in the courses obtained an average *High School Grade* of 82.9 which rises to an average 83.2 for students joining the experiment and 85.1 for students taking the test. About 50% of students in all three subsamples had studied in a Lyceum.²⁰ Both students joining the experiment and those who took the test expected to obtain an average grade of 25 in the intermediate test.²¹ In the final sample, about 43% are students of Microeconomics, 40% of Macroeconomics and 17% study Econometrics.

Students' answers to the on-line survey allowed us to build a self-reported measure of risk attitude. The question, the same as in the German Socio-Economic Panel (see Dohmen et al., 2011), is 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 them? Please tick a box on the scale, where 1 means: 'totally unwilling to take risks' and 10 means: 'very willing to take risks'." We built the variable *Risk Propensity*²² simply by reflecting these values chosen by students. *Risk Propensity* is on average 6.5 for the subsample of students joining the experiment, and almost the same for the students who effectively participated.²³

It would be interesting to analyze if risk preference has played a role in inducing students to join the experiment. Unfortunately, we have no measure of risk attitudes for those who chose not to be in the experimental exam and so we are not able to examine the issue of selection into the tournaments based on risk preferences. However, there is wide evidence in the literature showing that risk-averse people tend to avoid competitive settings (Eriksson et al., 2009; Dohmen and Falk, 2011).

Treatment groups were evenly balanced in the subsample of students joining the experiment (about 33% in each tournament) and remained so also in the subsample of students taking the intermediate test (about 34% in the *High* and *Medium Spread Tournament* and 32% in the *Low Spread Tournament*).

Tournament Performance represents the score obtained by students in the tournament (ranging from 0 to 10) and is our main dependent variable. Students obtained on average 4.1; the standard deviation is quite high (3.3), almost 20% of students obtain a score of zero and 32% fail to reach the threshold of 2 points required to obtain the bonus in the tournament part

¹⁹ Most of the existing laboratory studies focus on situations where the identity of the opponents is unknown. Mago et al. (2016), to the best of our knowledge, is the first paper investigating the effect of identification in contests. They show that knowing the opponent identity decreases wasteful effort in contests.

²⁰ In Italy, after lower secondary school pupils choose between a 'more academically oriented track' (Lyceum), or a more labor market-oriented track (Technical or Vocational). Students coming from more educated families typically choose a Lyceum, while those from poorer socio-economic backgrounds tend to enroll at vocational schools.

²¹ The *Expected Grade* turned out to be highly correlated to the effective grade obtained at the exam ($\rho = 0.49$; p -value = 0.00).

²² To elicit students' attitudes towards risk we also asked a second question regarding the preferred level of investment in a risky asset. Specifically: "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 of money permitting you, with the same probability, either to double the investment or to lose all the capital invested. Which amount would 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* ($\rho = 0.60$). In the following econometric analyses we use only *Risk Propensity* but very similar results are obtained using the second measure based on the amount to invest (estimates not reported but available upon request).

²³ Consistently with Dohmen et al. (2011) we find that the willingness to take risks is negatively correlated to student age and High School Grade. Women turn out to be significantly more risk-averse.

Table 1

Descriptive statistics. Mean and SD.

	Enrolled at the courses	Joining the experiment	Taking the exam	Difference A (p-value)	Difference B (p-value)	Difference C (p-value)
Female	0.488 (0.500)	0.487 (0.500)	0.522 (0.500)	−0.004 (0.945)	0.090 (0.055)	0.171 (0.007)
Age	22.236 (2.920)	21.757 (2.601)	21.788 (2.640)	−2.187 (0.000)	−1.185 (0.000)	0.151 (0.650)
High School Grade	82.934 (10.964)	83.249 (11.088)	85.110 (10.812)	1.437 (0.233)	5.754 (0.000)	9.136 (0.000)
Lyceum	0.492 (0.500)	0.503 (0.501)	0.492 (0.501)	0.050 (0.366)	−0.000 (0.998)	−0.054 (0.401)
Micro	0.417 (0.494)	0.471 (0.500)	0.432 (0.496)	0.244 (0.000)	0.038 (0.407)	−0.191 (0.003)
Macro	0.423 (0.423)	0.370 (0.484)	0.395 (0.490)	−0.243 (0.000)	−0.075 (0.108)	0.123 (0.047)
Econometrics	0.159 (0.159)	0.159 (0.366)	0.173 (0.379)	−0.002 (0.967)	0.036 (0.293)	0.069 (0.141)
Expected Grade		24.743 (2.419)	24.967 (2.287)			1.097 (0.000)
Risk Propensity		6.513 (1.808)	6.432 (1.840)			−0.399 (0.084)
High Spread Tournament		0.328 (0.470)	0.339 (0.474)			0.053 (0.377)
Medium Spread Tournament		0.339 (0.474)	0.342 (0.475)			0.018 (0.773)
Low Spread Tournament		0.333 (0.472)	0.319 (0.467)			−0.071 (0.242)
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: In the first three columns we report standard deviations in parentheses. In the last three columns we report the following differences: Difference A: Joining experiment–Not Joining; Difference B: Taking the exam–Not Taking the exam; Difference C: Taking the exam–Joining the Experiment but Not Taking the exam. In the last three columns we report in parentheses p-values for the test of equality of means.

of the exam. *Standard Grade* is the score students obtained in the piece rate part of the exam and ranges from 0 to 25. On average students score 12.5 points out of 25; about 25% obtain a score of 7 or less.^{24,25}

We are aware that the performance in the test can be considered a combination of preparedness (or long-term effort choice), test-taking ability and effort during the test, and student ability. Although we are unable to disentangle all these factors, thanks to random assignment, students should not differ, on average, in terms of ability or other predetermined characteristics, so we can identify the impact of different treatments on their effort alone.

For the tournament section of the exam, we also calculated how many points students lose for leaving the answers blank (*Skipped Questions*) or for mistakes (*Mistakes*). On average students lost 3.7 points for not attempting to answer the questions and 2.2 points for incorrect answers.

In the last three columns of [Table 1](#), we compare the three subsamples of students and find that those joining the experiment are significantly younger than non-participants. The sample of students actually taking the exam has a higher percentage of females, a higher *High School Grade* and higher expectations in the exam. We also find that students sitting the exam tend to be more risk averse. However, as shown in [Appendix B, Table B1](#), where we further investigate the determinants of the decision to take the exam through a linear probability model, the negative effect of risk propensity on the probability of taking the exam holds true only when we do not control for other individual characteristics. In these regressions (in some of which we also control for the rival's characteristics) we find that more able students and those who expect higher grades are more likely to participate, while the other controls do not exert a statistically significant effect.

It is worth highlighting the fact that the above-mentioned selection does not invalidate the design of our experiment, nor does it affect the results on incentive effects. To investigate the effect produced by prize spreads in tournaments separately from the effect due to selection into different tournaments, we need comparable individuals in the three treatment groups.

²⁴ If we consider the total score obtained in the intermediate test, given by the sum of the standard grade and the prize obtained in the competition part, 7% of students scored over 30 overall.

²⁵ Only 3 students took the midterm exam, passed the threshold and then chose to leave the experiment and take the standard exam.

Table 2

Students' characteristics across treatment groups.

Panel A. Students joining the experiment							
	(1) High Spread T.	(2) Medium Spread T.	(3) Low Spread T.	(4) (1) vs. (2) (<i>t</i> -stat)	(5) (1) vs. (3) (<i>t</i> -stat)	(6) (2) vs. (3) (<i>t</i> -stat)	(7) <i>F</i> -stat (<i>p</i> -value)
Female	0.508	0.453	0.500	0.054 (0.871)	0.08 (0.127)	−0.047 (−0.746)	0.44 (0.642)
Age	21.653	21.720	21.898	−0.067 (−0.211)	−0.245 (−0.742)	−0.177 (−0.530)	0.30 (0.744)
High School Grade	83.387	83.133	83.230	0.254 (0.184)	0.157 (0.110)	−0.097 (0.070)	0.02 (0.983)
Lyceum	0.548	0.492	0.468	0.056 (0.891)	0.080 (1.26)	0.024 (0.380)	0.84 (0.432)
Expected Grade	24.815	24.695	24.722	0.119 (0.365)	0.092 (0.328)	−0.027 (−0.088)	0.08 (0.920)
Risk Propensity	6.403	6.445	6.690	−0.042 (−0.188)	−0.287 (−1.25)	−0.245 (−1.06)	0.92 (0.397)
Obs.	124	128	126				
Panel B. Students taking the exam							
	(1) High Spread T.	(2) Medium Spread T.	(3) Low Spread T.	(4) (1) vs. (2) (<i>t</i> -stat)	(5) (1) vs. (3) (<i>t</i> -stat)	(6) (2) vs. (3) (<i>t</i> -stat)	(7) <i>F</i> -stat (<i>p</i> -value)
Female	0.549	0.505	0.510	0.044 (0.631)	0.039 (0.542)	−0.005 (−0.078)	0.23 (0.792)
Age	21.675	21.821	21.873	−0.146 (−0.400)	−0.197 (−0.534)	−0.051 (−0.133)	0.15 (0.861)
High School Grade	84.618	85.408	85.312	−0.790 (−0.525)	−0.695 (−0.438)	−0.095 (0.063)	0.16 (0.852)
Lyceum	0.539	0.485	0.448	0.054 (0.767)	0.091 (1.283)	0.037 (0.528)	0.83 (0.436)
Expected Grade	25	25.039	24.854	−0.039 (−0.116)	0.146 (0.468)	0.185 (0.571)	0.18 (0.838)
Risk Propensity	6.314	6.378	6.614	−0.065 (−0.260)	−0.301 (−1.127)	−0.236 (−0.892)	0.73 (0.485)
Obs.	102	103	96				

Notes: 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.

With this in mind, we initially assigned students joining the experiment to the three different treatments by creating groups that were comparable in terms of observable characteristics. The possibility of students switching to the standard exam after learning the assigned treatment could have invalidated our random assignment by introducing a self-selection element in the sample of students who effectively took the exam. Since only one part of the test was evaluated on the basis of student performance in the tournament, sorting on the basis of the assigned status was unlikely. In fact, both [Tables 1](#) and [B1](#) show that the decision to take the test was not affected by the type of tournament assigned. To further rule out this possibility, we also verify if groups are balanced for students who took the test.

In columns (1)–(3) of [Table 2](#) we report mean values for a number of individual characteristics for the three treatment groups in the subsample of initial participants (Panel A) and in the subsample of students actually took the test (Panel B). Columns (4)–(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 were unable to reject the hypothesis that the randomization was successful in creating comparable treatment groups as regards observable characteristics: there are no significant differences in terms of gender, *Age*, *High School Grade*, type of High School attended, *Expected Grade* and *Risk Propensity*. Our identifying assumption is that students in the three treatment groups also have on average similar unobservable characteristics.

Despite all our checks on the issue of selection, we acknowledge that, since students for ethical and administrative reasons were free to leave the experiment, we have a good deal of attrition (about 20% of students initially joining left). As is common in all randomized experiments, we are able to check only the balance on observable variables among treatment groups, the existence of a similar randomization on unobservable characteristics remains an assumption.

4. Students' performance and 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 tournament, participants have, on average, the same observable (and, by assumption, also unobservable) characteristics and thus we are able to exclude that selection effects are driving our results. This is true also

Table 3
Prize spreads and students' performance. OLS estimates.

	(1)	(2)	(3)	(4)	(5)	(6)
High Spread T.	−0.263 (0.458)	−0.332 (0.403)	−0.371 (0.406)	−0.380 (0.408)		
Medium Spread T.	−0.089 (0.459)	−0.186 (0.395)	−0.212 (0.397)	−0.235 (0.405)		
Spread					−0.093 (0.101)	−0.095 (0.102)
Macro	−0.211 (0.414)	−0.203 (0.378)	−0.150 (0.374)	−0.102 (0.400)	−0.149 (0.374)	−0.102 (0.399)
Econometrics	1.380*** (0.529)	1.397* (0.779)	1.400* (0.774)	1.580 (1.055)	1.399* (0.773)	1.580 (1.054)
Female		0.475 (0.336)	0.454 (0.336)	0.467 (0.346)	0.455 (0.335)	0.469 (0.345)
Age		−0.240** (0.100)	−0.252** (0.099)	−0.245** (0.100)	−0.252** (0.099)	−0.245** (0.100)
High School Grade		0.126*** (0.016)	0.124*** (0.016)	0.120*** (0.018)	0.123*** (0.016)	0.120*** (0.018)
Lyceum		0.501 (0.320)	0.529 (0.321)	0.553 (0.335)	0.529* (0.321)	0.553* (0.334)
Expected Grade		0.307*** (0.068)	0.318*** (0.068)	0.316*** (0.069)	0.318*** (0.068)	0.316*** (0.069)
Risk Propensity			−0.100 (0.096)	−0.103 (0.096)	−0.100 (0.096)	−0.103 (0.096)
Rival's Risk Propensity				−0.030 (0.086)		−0.030 (0.086)
Female Rival				−0.067 (0.333)		−0.065 (0.332)
Rival's Age				−0.030 (0.107)		−0.030 (0.107)
Rival's Lyceum				−0.336 (0.331)		−0.336 (0.331)
Absent Rival				−0.059 (0.467)		−0.056 (0.467)
Rival's Expected Grade				0.034 (0.075)		0.033 (0.074)
Constant	4.087*** (0.418)	−9.545*** (3.069)	−8.676*** (3.187)	−8.257* (4.781)	−8.499*** (3.208)	−8.066* (4.796)
Observations	301	301	301	301	301	301
Adjusted R-squared	0.018	0.293	0.294	0.283	0.296	0.285

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% level, respectively.

for students effectively taking the test, since as we have shown in Tables 2 and B1, selection into participation does not depend on treatment regimes.

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

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

where the dependent variable, *Tournament Performance_i*, is the total score that the student obtains in the tournament 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 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 predetermined characteristics and cognitive abilities and β_3 the vector of their coefficients; ε_i is an error term.

Table 3 reports OLS estimates for the impact of prize 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 produces any effect on students' performance: the coefficients are negative but far from being statistically significant. These results are in contrast with the standard theory of tournaments, but consistent with findings of some field experiments (such as Leuven et al., 2011) and lab experiments reported in the literature review section.

The second specification adds among controls student gender, age, high school grade, type of high school and expected grade. Older students achieve worse grades while the measures of ability tend to affect the score obtained at the competition positively (in particular, 10 points more of *High School Grade* increase by 1.26 the performance in the competition, an effect that corresponds to about 0.4 standard deviations of the dependent variable; instead, the expectation to obtain a grade one point higher increases tournament performance by about 0.3 points). The impact of prize spreads is still statistically not significant.

In the third specification we also control for students' self-reported risk attitude (*Risk Propensity*). Adding this control does not change our main results: taking part in tournaments with larger prizes does not affect performance with respect to the tournament with the lowest spread. The attitude towards risk does not seem to affect student performance directly.

Since individuals in tournaments take strategic decisions that could be affected by the characteristics of the competitors, in column (4) we control for some opponents' characteristics: gender, *Age*, *Lyceum*, *Expected Grade*, *Risk Propensity* and presence at the test.²⁶ None of the competitor's characteristics turn out to be statistically significant and in addition the impact of tournament type is still not significant.

We do not find statistically significant differences between men and women as regards prize spreads, although, women seem to perform worse in *High Spread Tournaments* but better in *Medium Spread Tournaments*.

We have also investigated if the prize spread has a different impact on performance depending on other individual characteristics, using interactions of the prize spread with characteristics, such as age or High School Grade, but find that these interactions are not statistically significant.

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 (−0.093) but not significant impact on students' performance in the tournament (columns 5 and 6 of Table 3).

We also try to understand whether prize spreads do not affect effort decisions because the students obtain utility from the mere fact of outperforming their opponent (the so-called “Joy of winning”, Sheremeta, 2010). Although we do not have a standard measure of “joy of winning”, we consider the following question in our on-line survey: “Generally, when I participate in a competition: a) I work hard because it's important for me to win; b) I put in a fair amount of effort because for me the important thing is to participate; c) I do not work much because I think I cannot win”. Using the answers to this question we build a dummy variable, *Joy of Winning*, taking the value of one for students choosing option a) and zero otherwise. Our results show that *Joy of Winning* has a positive but not statistically significant impact on students' performance. However, the inclusion of this variable among control variables does not affect our main results (estimates not reported to save space).

Finally, in order to understand whether tournaments have an incentive effect different from standard individual incentives we compare student performance in the piece-rate and tournament parts of the exam. We measure students' performance in the two parts of the exam as the percentage of total points available (*Grade*) gained in each part and create two observations for each student (tournament and piece-rate). We estimate with OLS the impact of tournaments and treatment status using as a dependent variable *Grade* and including students' fixed effects. Our findings show that students perform significantly worse in tournaments (−0.870) and the effect is magnified for risk-averse students (−1.468). In addition, the negative effect is smaller in magnitude in competitions with the low and medium spread (about −0.600), while it increases in competitions with the highest spread (−1.281) especially for risk-averse students (results not reported and available upon request). However, these findings represent only suggestive evidence, under the assumption that the two parts of the exam are similar in the level of difficulty. In fact, in order to test this effect properly it would have been necessary to randomly assign different parts of the program to the tournament part and to the piece-rate part but our experiment was not designed to pursue this aim.

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.

In order to empirically analyze this issue, we use the measure of students' risk-aversion elicited in the preliminary survey and 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 4. In column (1) we only control for dummies of the course attended. In column (2) we control also for *Female* and *Age*. In column (3) we control for three measures of student ability (*High School Grade*, *Lyceum* and *Expected Grade*) while in column (4) we include competitor characteristics (*Gender*, *Age*, *Lyceum*, *Expected Grade*, *Risk Propensity*, *Absent*). The coefficients of control variables are not reported to save space.

Interestingly, we find that the effect of prize spreads on student performance depends on their propensity to take risks. Results are very similar in all the specifications. For students highly averse to risk (*Risk Propensity* equal to 1) we find a

²⁶ Given that pairs of students were formed on the basis of their High School Grades, the correlation between their own and Rival's High School Grade is 0.98. For this reason we do not control for rival's characteristics. If we include Rival's High School Grade the results are very similar but the two High School Grade variables are not significant because of their high collinearity.

Table 4

Risk aversion and students' performance in tournaments. OLS estimates.

	(1)	(2)	(3)	(4)	(5)	(6)
High Spread T.	−3.220** (1.541)	−3.168** (1.500)	−2.974** (1.376)	−2.862** (1.411)	−1.708** (0.755)	−1.671** (0.764)
Medium Spread T.	−3.396** (1.570)	−2.968* (1.560)	−2.886** (1.411)	−2.875** (1.429)	−1.034 (0.787)	−1.007 (0.797)
High Spread T.* Risk Propensity	0.448* (0.231)	0.421* (0.227)	0.400** (0.199)	0.381* (0.204)		
Medium Spread T.* Risk Propensity	0.503** (0.233)	0.430* (0.231)	0.409** (0.203)	0.404** (0.205)		
Risk Propensity	−0.426*** (0.146)	−0.407*** (0.143)	−0.359*** (0.135)	−0.353*** (0.135)		
Risk Taker					−1.553** (0.713)	−1.589** (0.721)
High Spread T.* Risk Taker					1.942** (0.890)	1.951** (0.911)
Medium Spread T.* Risk Taker					1.160 (0.908)	1.124 (0.919)
Course dummies	Yes	Yes	Yes	Yes	Yes	Yes
Individual characteristics	No	Yes	Yes	Yes	Yes	Yes
Ability (High School Grade; Lyceum; Expected Grade)	No	No	Yes	Yes	Yes	Yes
Competitor characteristics	No	No	No	Yes	No	Yes
Observations	301	301	301	301	301	301
Adjusted R-squared	0.030	0.088	0.301	0.289	0.303	0.296

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% level, respectively.

very strong negative effect both for *High Spread Tournament* and for *Medium Spread Tournament*: students' performance is about 2.5 points lower when assigned to these two types of tournaments (for example, in column 4 is -2.48 for *High Spread Tournament* and -2.47 for *Medium Spread Tournament*, both statistically significant at the 5% level). The magnitude of these effects is quite marked: it corresponds to about 0.75 standard deviations of the dependent variable.

However, the interaction terms *High Spread Tournament**(*Risk Propensity*) and *Medium Spread Tournament**(*Risk Propensity*) are both positive (around 0.4 points) and statistically significant, implying that for an individual with an average risk propensity of 6 the impact of both treatment groups is negative but not statistically different from zero. For example, in column (4) we show that for an individual with average risk-aversion the impact of being assigned to the *High Spread Tournament* is -0.577 ($= -2.862 + (6 \cdot 0.381)$) (t -stat $= -1.35$).²⁷

It is worthwhile to notice that, since we do not have information on the risk propensity of individuals who decided to not join the experiment, we cannot exclude that these individuals are more risk prone, since joining could be considered a safer move.²⁸ On the other hand, a factor that could counterbalance this effect is the tendency of risk-averse individuals to avoid competitive settings (Eriksson et al., 2009; Dohmen and Falk, 2011). If the former factor prevails, the effects of risk preferences on performance could be exacerbated by having more risk-averse people in the pool taking the exam. However, the proportion of students not joining the experiment is not so high, only 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 4 we replicate specifications (3) and (4) but use a dummy variable *Risk Taker*, taking the value of one for values of *Risk Propensity* higher than 6 (and zero otherwise), and interact *Risk Taker* with our two treatment dummies. We find a negative and statistically significant effect of competing in *High Spread Tournament* for students who are risk-averse (-1.671 , corresponding to about 0.5 standard deviation of the dependent variable), significant at the 5% level, while this effect approaches zero ($0.279 = -1.671 + 1.951$) with a t -stat of 0.57 for individuals more prone to taking risks. Being assigned to the *Medium Spread Tournament* produces a negative but not statistically significant effect (-1.007 , t -stat $= -1.26$) for risk-averse individuals. The impact of *Medium Spread Tournament* shrinks again to almost zero ($0.117 = -1.007 + 1.124$; t -stat $= 0.25$) for more risk-prone students.

Finally, to check the robustness of our findings in Table 5, we replicate the estimates of Table 4 but use *Spread* in linear form (instead of using the two dummies for the two treatment regimes) and interact it with *Risk Propensity*. We find that one point of spread decreases performance of about 0.7 (significant at the 5% level) for risk-averse individuals, whereas its effect is close to zero for more risk prone individuals (for example, in column 4 for an individual with average risk aversion the impact is $-0.138 = -0.718 + 0.097 \cdot 6$, t -stat $= -1.30$).²⁹ Qualitatively similar results are found with the dummy *Risk Taker*.

²⁷ The coefficient of *Risk Propensity* in Table 4 suggests that as risk aversion reduces, the performance of participants tend to be lower.

²⁸ This could actually be the case as in our setting students can always drop out of the experiment and sit the standard exam at the end of the semester (even if they failed the intermediate exam).

²⁹ Results shown in Tables 4 and 5 do not change qualitatively if we control for our measure of *Joy of Winning*.

Table 5

Risk aversion and students' performance in tournaments. Spread in linear form. OLS estimates.

	(1)	(2)	(3)	(4)	(5)	(6)
Spread	−0.814** (0.387)	−0.626* (0.350)	−0.747** (0.345)	−0.718** (0.353)	−0.423** (0.186)	−0.411** (0.189)
Risk Propensity	−0.573** (0.239)	−0.370* (0.213)	−0.497** (0.212)	−0.482** (0.215)		
Spread*Risk Propensity	0.114* (0.058)	0.087* (0.051)	0.101** (0.050)	0.097* (0.051)		
Risk Taker					−2.441** (1.022)	−2.381** (1.041)
Spread*Risk Taker					0.481** (0.220)	0.462** (0.225)
Course dummies	Yes	Yes	Yes	Yes	Yes	Yes
Individual characteristics	No	Yes	Yes	Yes	Yes	Yes
Ability	No	No	Yes	Yes	Yes	Yes
Rival's characteristics	No	No	No	Yes	No	Yes
Observations	301	301	301	301	301	301
Adjusted R-squared	0.031	0.252	0.303	0.291	0.307	0.296

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% level, respectively.

Table 6

The probability of failing to reach the threshold at the competition part. Linear Probability Model.

	(1)	(2)	(3)	(4)	(5)	(6)
High Spread T.	0.026 (0.061)	0.041 (0.061)	0.182 (0.218)	0.143 (0.222)	0.176* (0.104)	0.181* (0.108)
Medium Spread T.	0.018 (0.058)	0.028 (0.059)	0.257 (0.204)	0.238 (0.206)	0.185* (0.103)	0.200* (0.106)
Risk Propensity	0.001 (0.015)	−0.000 (0.015)	0.020 (0.018)	0.015 (0.019)		
High Spread T.* Risk Propensity			−0.024 (0.032)	−0.015 (0.033)		
Medium Spread T.* Risk Propensity			−0.037 (0.031)	−0.032 (0.030)		
Risk Taker					0.182** (0.089)	0.184** (0.092)
Risk Taker*High Spread					−0.212* (0.127)	−0.206 (0.130)
Risk Taker*Medium Spread					−0.235* (0.126)	−0.246* (0.127)
Course dummies	Yes	Yes	Yes	Yes	Yes	Yes
Individual characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Ability	Yes	Yes	Yes	Yes	Yes	Yes
Rival's characteristics	No	Yes	No	Yes	No	Yes
Observations	301	301	301	301	301	301
Adjusted R-squared	0.192	0.191	0.190	0.188	0.197	0.189

Notes: The dependent variable is the dummy *Failing*. 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% level, respectively.

All these findings suggest that risk-averse individuals tend to perform worse in tournaments with large spreads. Since in our setting individuals have to take the piece rate and the tournament sections of our exam jointly, it is likely that risk-averse students who dislike the high uncertainty in rewards embedded in large spread tournaments tend to save the cost related to effort.³⁰

However, whether this choice corresponds to effectively opting out of the tournament is difficult to say. To better investigate this aspect, we look at students (32%) who failed to reach the 20% threshold for the tournament part of the exam (less than 2 out of 10 points). Considering students' achievement in the exam, we created a dummy variable *Failing* taking the value of one for students who did not reach the minimum threshold for obtaining the prize in the competition part.

We estimate a Linear Probability Model for the probability of failing to reach the threshold on the basis of treatment status and risk propensity. Results are reported in Table 6. In the first two columns (without using interactions) we do not find any effect of the type of tournament on the probability of failing. *Risk Propensity* does not influence this probability.

However, when we interact *Risk Propensity* with treatment status, and especially when we interact the dummy *Risk Taker* with the treatment status (columns 5 and 6), we find that risk-averse individuals tend to fail more often in *High* and *Medium Spread Tournaments* (significant at the 10% level).

On the other hand, risk-takers are not affected by the type of tournament as regards the probability of failing (for example, for *High Spread Tournaments*, $0.181 - 0.206 = -0.025$, not statistically significant).

Similar results are found when we consider as dependent variable *Skipped Questions* (points lost due to students not attempting the questions). If individuals have chosen to not participate in the competition we should observe that the skipped questions in the tournament part of the test are more frequent for risk-averse individuals in large spread tournaments. As expected, we find that in high and medium spread tournaments students with high risk-aversion tend to leave the answers blank (2–3 more points missed than in the *Low Spread Tournament*), while more risk-prone individuals miss fewer points (results available upon request). Although the statistical significance of many coefficients is not high, the qualitative findings are pretty clear: by skipping a higher number of questions highly risk averse students choose to spend less effort. However, it is difficult to claim that they have given up completely since in that case they would have skipped all the questions included in the tournament part of the exam.³¹

All in all, our findings consistently suggest that risk-averse individuals tend to perform worse in tournaments with large spreads. These results are not the outcome of selection, since – as we have shown – individuals of comparable ability were competing in each type of tournament. On the other hand, our results suggest that risk-averse individuals might have left the tournament had this option been available. Since in our setting individuals have to take both the piece rate and tournament sections of the exam, risk-averse students who dislike the high uncertainty in rewards embedded in large spread tournaments have chosen to save on the provision of effort in the tournament stage. However, they still came to the test because of the presence of the piece-rate section that accounts for 75% of the exam. Even if our evidence is not strong enough for clear predictions, it could be that risk-averse agents would simply have chosen not to participate in the tournament were this possible.

Our finding that larger spreads might decrease performance are in contrast with the standard version of tournament theory that, by assuming typically risk-neutral agents, claims that the agents' effort level increases with the prize spread. However, our empirical findings can be reconciled with the theoretical predictions considering that – as shown by Nalebuff and Stiglitz (1983) – when agents are risk-averse their utility can decrease in large spread tournaments. In addition, as explained by Bull et al. (1987): “the agents will have to take into account two sources of uncertainty. ... the distribution of the prizes induced by the randomness in production [...] and the uncertainty concerning how the specific tournament that the agent enters will be played”. The many sources of uncertainty (the randomness in obtaining the prize, the bounded ability of understanding all the consequences of the game they are involved in, the difficulty in forecasting how the competitors will play) are amplified as the spread increases and this could have induced the students participating in our experiment to consider high spread tournaments as a sort of lottery in which they had little desire to invest much effort. Finally, Sheremeta and Wu (2012) show that when the agents' risk aversion is accompanied by the hypothesis of a utility function that is not additively separable in income and effort, an increase in the prize spread can effectively worsen performance to save on the disutility of effort.

The policy implications of our results are important. In real labor markets in many situations there is no perfect mobility of workers (due to search and matching frictions or specific investments) or simply labor markets are not tight and it is not easy to find a new job.³² In these cases, after the introduction of an incentive system as a high spread tournament, agents might not find it easy to move to another job (at least, not in the short-run) and thus they might decide to spend little effort. Our findings suggest that when individuals cannot easily select in and out of a tournament, high spreads can be counterproductive. This is especially the case when dealing with risk averse agents.

6. The performance of students under a piece-rate incentive system

In principle, the assignment of students to the three different tournaments might have affected their performance not only in the competitive 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 run a number of regressions considering as a dependent variable *Standard Grade*. In all specifications in Table 7 we do not find any statistically significant treatment effect on students' performance in the piece rate section, neither when we use the treatment dummies alone, nor when we interact them with *Risk Propensity*.

³⁰ In fact, we do not find that more risk-averse students tend to leave the experiment if assigned to tournaments with larger spreads.

³¹ We have also considered as outcome variable the points lost due to incorrect answers (*Mistakes*). We do not find any significant effect either for the type of tournament or for the degree of risk aversion.

³² Alternatively, the agent could be interested in the job for many other reasons (interest in the work, location, relationships with colleagues, etc.) beyond the incentive system.

Table 7

The impact of prize spreads on standard grade. OLS estimates.

	(1)	(2)	(3)	(4)	(5)	(6)
High Spread T.	0.808 (1.054)	0.473 (0.913)	0.310 (0.918)	−0.262 (3.493)		
Medium Spread T.	−0.327 (1.020)	−0.664 (0.922)	−0.778 (0.939)	−4.713 (3.022)		
Spread					0.082 (0.230)	−0.033 (0.865)
Risk Propensity		−0.587*** (0.214)	−0.597*** (0.213)	−0.813** (0.326)	−0.593*** (0.213)	−0.663 (0.529)
High Spread T.*Risk propensity				0.078 (0.518)		
Medium Spread T.* Risk Propensity				0.607 (0.449)		
Spread* Risk Propensity						0.018 (0.128)
Course dummies	Yes	Yes	Yes	Yes	Yes	Yes
Individual characteristics	No	Yes	Yes	Yes	Yes	Yes
Ability	No	Yes	Yes	Yes	Yes	Yes
Rival's characteristics	No	No	Yes	Yes	Yes	Yes
Observations	301	301	301	301	301	301
Adjusted R-squared	0.002	0.235	0.228	0.228	0.227	0.225

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% level, respectively.

Neither do we find any effect when (columns 5 and 6) the variable *Spread* is used. These results suggest that students have not been affected by prize spreads when deciding how hard to study for the piece rate part of the test, and there were no spillovers from one part of the exam to the other.

Since students' performances in 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 (not reported) remain qualitatively unchanged with respect to the estimates shown in the paper.

7. Concluding remarks

We have run a field experiment to investigate the incentive effect of higher prize spreads in tournaments and analyze whether this effect is heterogeneous depending on individual risk aversion. Our experiment has involved about 300 students competing in pairs to obtain the best performance in rank-order tournaments with different prize spreads. These students were analyzed in their natural environment, in a setting in which tournaments should be the ideal incentive scheme as the uncertainty deriving from common shocks (such as teaching material, teaching quality, difficulty of test questions) is probably more important than the uncertainty deriving from agents' idiosyncratic shocks (shocks during exam preparation, mood on exam day, etc.). However, since students for ethical and administrative reasons were free to leave the experiment, we had a good deal of attrition (about 20% of students initially joining left). Although analogously to all randomized experiments we are able to check only the balance on observable variables among treatment groups – and the existence of a similar randomization on unobservable characteristics remains an assumption – from some of our checks a significant selection of students according to their degree of risk aversion does not emerge.

We have found that – in contrast with standard theoretical predictions – agents do not make more effort when prize spreads are wider. This finding is in line with the evidence provided by [Leuven et al. \(2011\)](#) who also show that tournaments have no incentive effects.

The absence of incentive effects is also in contrast with results found by many laboratory experiments. As argued by [Leuven et al. \(2011\)](#), this might depend on the fact that while in lab experiments participants can do nothing besides perform the task, in field experiments they can use their time in other ways.

Since in our experiment the probability of winning in each competition was 50% (considering that the two agents in competition were of similar ability), we can rule out the absence of incentive effects due to low expected payoffs. In addition, in the treatment with the highest spread, the winning student could score 8 points, which is quite a sizeable gain as the maximum grade at the exam was 30.

It could also be that students in our sample are not interested in the size of the spread either because they care about winning per se (“joy of winning”) or because they have social preferences that counter-balance the benefits deriving from an increase in the spread. In our framework, in fact, the increase in the spread leads to a reduction in the loser's prize and in the case of altruistic or inequity-averse agents this might generate disutility. While we do not have measures of social

preferences, we have tried to investigate the role played by “the joy of winning” but have not found a statistically significant effect.

Explanations based on the joy of winning or social preferences are not fully consistent with another crucial finding of our analysis, that is that while the effect of large spreads is not significant for risk-takers, it is clearly detrimental to the performance of risk-averse individuals.

These findings support the view that risk averse individuals tend to avoid situations that involve a high level of uncertainty and when the “exit option” is not feasible they tend to exert little effort. An increase in prize spreads implies both a raise in the uncertainty coming from the increase in the variance of outcomes and an increase in the uncertainty regarding strategies adopted by other agents. It is likely that these sources of uncertainty have weakened the relationship between effort and rewards and have discouraged risk-averse students assigned to high-spread tournaments from exerting much effort in the tournament section of the exam.

The policy implications of our findings – showing essentially that risk-averse agents do not like to be involved in high-stakes tournaments – are quite important, since in many contexts the mobility of workers is not perfect (due for example to search and matching frictions or slackness of local labor markets). In a world without frictions, agents could self-select into incentive schemes according to their characteristics and therefore risk-averse agents would avoid jobs in which compensation is based on a tournament scheme. However, when agents cannot easily and quickly move to another job, our evidence implies that the introduction of a tournament could have detrimental effects on the incentives of many individuals.

Appendix A. Instructions to students

In this academic year the students of this course can take the exam in the standard way or, alternatively, by joining an experiment we are running, they can take the exam based on an alternative scheme, in which about 25% of the grade is assigned on the basis of the relative performance of two competing students.

The alternative scheme consists of an intermediate and a final test. The intermediate test will cover the first part of the class program and will take place on 18th April. Students who pass the intermediate test will take a final test at the end of the course that covers the second part of the program. The final grade at the exam will be computed as the average of the grades obtained at the intermediate test and at final test. Students who sit the intermediate test – if unsatisfied – can decide to sit the standard exam at the end of the course.

The intermediate test will consist in two parts: one (individual part) evaluated on the basis of the student's absolute performance and the other (competition part) evaluated according to the student's performance with respect to a randomly chosen colleague.

Individual part:

1. The maximum score that can be obtained in the individual part of the intermediate test is 25.
2. The test consists of 12 questions/exercises. The score simply depends on the number of correct answers provided.
3. The individual section will cover the following chapters of your textbook....

Competition part:

1. Each student is paired (using a random procedure) with another student of similar expected ability.³³
2. Pairs of students will be randomly assigned to three types of competition: A, B, C.
 - Type A: The student who gets the highest score within each pair will receive an 8 point bonus; the student who gets the lowest score will receive a 2 point bonus.
 - Type B: The student who gets the highest score within each pair will receive a 7 point bonus; the student who gets the lowest score will receive a 3 point bonus.
 - Type C: The student who gets the highest score within each pair will receive a 6 point bonus; the student who gets the lowest score will receive a 4 point bonus.
3. The score obtained at the individual part will be added to the bonus earned at the competition part to determine the grade of the intermediate test.
4. The test consists of 6 questions or exercises and you may get a maximum score of 10.
5. In each of the competitions, the bonus will be awarded only if the student has reached a minimum 20% of correct answers in the competition part.
6. No draws are allowed. In the case of assignments with the same score in a pair, the teacher will carefully evaluate each aspect of the tasks in order to select a single winner. In case of identical performance, the winner will be randomly selected.
7. If one of the students in a pair does not show up for the exam, the competitor will be randomly paired with another student with the same expected ability and the same type of tournament in order to determine the tournament bonus.
8. The competition part will cover the following chapters...

³³ In case an odd number of students joining the experiment, we plan to randomly selected one student and evaluate his/her performance with respect to the performance of another student already in a pair (without any consequence for the outcome of the latter).

The final test will consist of questions and exercises covering all the remaining teaching material and will allow students to score a maximum of 30 points on the basis of their absolute performance.

The time available for completing each test (intermediate and final) is 105 min.

The pairs of competing students and the type of competition to which they are assigned will be published on the course webpage.

To join the experiment students have one week to complete an on-line form and a short questionnaire.

Appendix B. Probability of taking the test

As a further check aimed at ensuring 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 for the test on the sample of students joining the experiment (378). Estimation results from a Linear Probability Model are reported in Table B1.

In the first specification, we only control for course dummies and types of tournament. In the second column we add some individual characteristics (gender and age) and cognitive abilities (measured by the typology of high school attended, the *High School Grade* and the *Expected Grade*). In the third column we add *Risk Propensity*. In column (4) we add some of the rival's characteristics (gender, age, lyceum, expected grade, absent, risk propensity). In column (5) we use the dummy *Risk Taker* instead of *Risk Propensity*.

Table B1

Determinants of students' decision to take the exam. Linear Probability Model.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High Spread T.	0.060 (0.051)	0.055 (0.049)	0.049 (0.049)	0.059 (0.050)	0.058 (0.050)	0.054 (0.051)	0.054 (0.051)	0.050 (0.049)
Medium Spread T.	0.042 (0.051)	0.043 (0.048)	0.038 (0.048)	0.048 (0.048)	0.048 (0.048)	0.037 (0.051)	0.042 (0.051)	0.038 (0.048)
Macro	0.119*** (0.045)	0.102** (0.046)	0.111** (0.046)	0.111** (0.049)	0.113** (0.049)	0.126*** (0.045)	0.113** (0.045)	0.104** (0.046)
Econometrics	0.136** (0.055)	0.095 (0.100)	0.097 (0.098)	0.118 (0.131)	0.119 (0.132)	0.124** (0.056)	0.107* (0.056)	0.118 (0.098)
Female		0.023 (0.041)	0.020 (0.041)	0.018 (0.040)	0.018 (0.040)		0.087** (0.042)	0.022 (0.041)
Age		−0.014 (0.014)	−0.015 (0.013)	−0.014 (0.014)	−0.015 (0.014)			−0.018 (0.014)
High School Grade		0.010*** (0.002)	0.010*** (0.002)	0.010*** (0.002)	0.010*** (0.002)			0.011*** (0.002)
Lyceum		−0.024 (0.039)	−0.018 (0.040)	−0.013 (0.040)	−0.017 (0.040)			
Expected Grade		0.020** (0.009)	0.022** (0.009)	0.023** (0.009)	0.022** (0.009)			
Risk Propensity			−0.016 (0.011)	−0.016 (0.012)		−0.020* (0.011)	−0.017 (0.011)	−0.012 (0.011)
Absent Rival				0.073 (0.058)	0.078 (0.058)			
Rival's Risk Propensity				0.018 (0.011)				
Female Rival				0.047 (0.040)	0.048 (0.041)			
Rival's Age				−0.004 (0.013)	−0.003 (0.013)			
Rival's Lyceum				−0.028 (0.041)	−0.024 (0.041)			
Rival's Expected Grade				0.010 (0.010)	0.011 (0.010)			
Risk Taker					−0.058 (0.043)			
Risk Taker Rival					0.052 (0.044)			
Constant	0.697*** (0.046)	−0.365 (0.418)	−0.240 (0.421)	−0.609 (0.623)	−0.593 (0.613)	0.828*** (0.088)	0.776*** (0.092)	0.250 (0.354)
Observations	378	378	378	378	378	378	378	378
Adjusted R-squared	0.017	0.123	0.125	0.130	0.126	0.022	0.031	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% level, respectively.

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 hence 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, in the 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.

Finally, in the last three columns (6–8) we estimate some specifications with the aim of better investigating the relationship between risk propensity and the probability of taking the exam. While it is true that considering only the mean values of risk-propensity it seems that there is some selection of more risk-averse subjects into taking the exam, we show in column (6) that the negative effect of risk propensity on the probability of taking the exam holds true only when we do not control for other individual characteristics. This is probably due to the fact that risk propensity is correlated with these characteristics. For instance, a number of papers show that women are more risk averse than men (Croson and Gneezy, 2009; Dohmen et al., 2011; De Paola, 2012).

As shown in column (7) of Table B1 the negative effect of risk propensity on the probability of sitting the exam becomes statistically not significant when we include among controls the dummy *Female*. In addition, some works show an inverse relationship between cognitive abilities and risk aversion (Dohmen et al., 2010; De Paola, 2012). When we include measures of ability as explanatory variables (col. 8), the impact of risk propensity shrinks and becomes not statistically significant. In general, students' decision to show up for 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.

References

- Abrevaya, J., 2002. Ladder tournaments and underdogs: Lessons from professional bowling. *J. Econ. Behav. Org.* 47 (1), 87–101.
- Anderson, L.R., Freeborn, B.A., 2010. Varying the intensity of competition in a multiple prize rent seeking experiment. *Public Choice* 143 (1–2), 237–254.
- Araujo, F.A., Carbone, E., Conell-Price, L., 2016. The slider task: an example of restricted inference on incentive effects. *J. Econ. Sci. Assoc.* 2 (1), 1–12.
- Ariely, D., Bracha, A., Meier, S., 2009. Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *Am. Econ. Rev.* 99 (1), 544–555.
- Bandiera, O., Barankay, I., Rasul, I., 2005. Social preferences and the response to incentives: evidence from personnel data. *Q. J. Econ.* 120 (3), 917–962.
- Baye, M.R., Kovenock, D., De Vries, C.G., 1996. The all-pay auction with complete information. *Econ. Theory* 8 (2), 291–305.
- Becker, B.E., Huselid, M.A., 1992. The incentive effects of tournament compensation systems. *Adm. Sci. Q.* 37 (2), 336–350.
- Bull, C., Schotter, A., Weigelt, K., 1987. Tournaments and piece rates: an experimental study. *J. Polit. Econ.* 95 (1), 1–33.
- Buser, T., Niederle, M., Oosterbeek, H., 2014. Gender, competitiveness and career choices. *Q. J. Econ.* 129 (3), 1409–1447.
- Cason, T.N., Masters, W.A., Sheremeta, R.M., 2010. Entry into winner-take-all and proportional-prize contests: an experimental study. *J. Public Econ.* 94 (9–10), 604–611.
- Charness, G., Kuhn, P., 2011. Lab labor: What can labor economists learn from the lab? *Handbook Labor Econ.* 4, 229–330.
- Cornes, R., Hartley, R., 2012. Risk aversion in symmetric and asymmetric contests. *Econ. Theory* 51 (2), 247–275.
- Croson, R., Gneezy, U., 2009. Gender differences in preference. *J. Econ. Lit.* 47, 448–474.
- De Paola, M., 2012. The determinants of risk aversion: the role of intergenerational transmission. *Ger. Econ. Rev.* 14 (2), 214–234.
- De Paola, M., Scoppa, V., 2011. Frequency of examinations and student achievement in a randomized experiment. *Econ. Educ. Rev.* 30, 1416–1429.
- De Paola, M., Scoppa, V., 2015. Procrastination, academic success and the effectiveness of a remedial program. *J. Econ. Behav. Org.* 115, 217–236.
- De Paola, M., Scoppa, V., Nisticò, R., 2012. Monetary incentives and student achievement in a depressed labor market: results from a randomized experiment. *J. Hum. Capital* 6 (1), 56–85.
- Dechenaux, E., Kovenock, D., Sheremeta, R.M., 2015. A survey of experimental research on contests, all-pay auctions and tournaments. *Exp. Econ.* 18 (4), 609–669.
- Delfgaauw, J., Dur, R., Sol, J., Verbeke, W., 2013. Tournament incentives in the field: gender differences in the workplace. *J. Labor Econ.* 31 (2), 305–326.
- DellaVigna, S., Pope, D., 2017. What motivates effort? Evidence and expert forecasts. *Rev. Econ. Stud.* 1–41.
- Di Pietro, G., 2013. Military conscription and university enrolment: evidence from Italy. *J. Popul. Econ.* 26 (2), 619–644.
- Dohmen, T., Falk, A., 2011. Performance pay and multidimensional sorting: productivity, preferences, and gender. *Am. Econ. Rev.* 101, 556–590.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., 2010. Are risk aversion and impatience related to cognitive ability? *Am. Econ. Rev.* 120, 256–271.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., Wagner, G.G., 2011. Individual risk attitudes: measurement, determinants, and behavioral consequences. *J. Eur. Econ. Assoc.* 9 (3), 522–550.
- Drugov, M., Ryvkin, D., 2017. Biased contests for symmetric players. *Games Econ. Behav.* 103, 116–144.
- Ehrenberg, R.G., Bognanno, M.L., 1990. Do tournaments have incentive effects. *J. Polit. Econ.* 98 (6), 1307–1324.
- Eriksson, T., 1999. Executive compensation and tournament theory: empirical tests on Danish data. *J. Labor Econ.* 17 (2), 262–280.
- Eriksson, T., Teyssier, S., Villeval, M.C., 2009. Self-selection and the efficiency of tournaments. *Econ. Inq.* 47 (3), 530–548.
- Erkal, N., Gangadharan, L., Koh, B.H., 2018. Monetary and non-monetary incentives in real-effort tournaments. *Eur. Econ. Rev.* 101, 528–545.
- Falk, A., Fehr, E., Huffman, D., 2008. The Power and Limits of Tournament Incentives. Mimeo.
- Fershtman, C., Gneezy, U., 2011. The trade-off between performance and quitting in high-power tournaments. *J. Eur. Econ. Assoc.* 9, 318–336.
- Gill, D., Prowse, V., 2012. A structural analysis of disappointment aversion in a real effort competition. *Am. Econ. Rev.* 469–503.
- Gneezy, U., Rustichini, A., 2000. Pay enough or don't pay at all. *Q. J. Econ.* 115, 791–810.
- Green, J., Stokey, N., 1983. A Comparison of Tournaments and Contracts. *J. Polit. Econ.* 91 (3), 349–364.
- Harbring, C., Irlenbusch, B., 2003. An experimental study on tournament design. *Labour Econ.* 10 (4), 443–464.
- Harbring, C., Irlenbusch, B., 2008. How many winners are good to have?: On tournaments with sabotage. *J. Econ. Behav. Org.* 65 (3), 682–702.
- Harbring, C., Lünser, G.K., 2008. On the competition of asymmetric agents. *Ger. Econ. Rev.* 9 (3), 373–395.
- Hillman, A.L., Katz, E., 1984. Risk-averse rent seekers and the social cost of monopoly, power. *Econ. J.* 94, 104–110.
- Hong, F., Hossain, T., List, J.A., 2015. Framing manipulations in contests: a natural field experiment. *J. Econ. Behav. Org.* 118, 372–382.
- Jenkins, G.D., Mitra, A., Gupta, N., Shaw, J., 1998. Are financial incentives related to performance? A meta-analytic review of empirical research. *J. Appl. Psychol.* 83 (5), 777–787.
- Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. *Econometrica* 47 (2), 263–291.
- Kräkel, M., Nieken, P., 2015. Relative performance pay in the shadow of crisis. *Eur. Econ. Rev.* 74, 244–268.
- Lazear, E.P., Rosen, S., 1981. Rank-order tournaments as optimum labor contracts. *J. Polit. Econ.* 89 (5), 841–864.

- Leuven, E., Oosterbeek, H., Sonnemans, J., Van Der Klaauw, B., 2011. Incentives versus sorting in tournaments: evidence from a field experiment. *J. Labor Econ.* 29 (3), 637–658.
- Mago, S., Samak, A., Sheremeta, R., 2016. Facing your opponents: social identification and information feedback in contests. *J. Conflict Resolut.* 60 (3), 459–481.
- Millner, E.L., Pratt, M.D., 1991. Risk aversion and rent-seeking: An extension and some experimental evidence. *Public Choice* 69 (1), 81–92.
- Mohnen, A., Pokorny, K., Sliwka, D., 2008. Transparency, inequity aversion, and the dynamics of peer pressure in teams: theory and evidence. *J. Labor Econ.* 26 (4), 693–720.
- Nalebuff, B., Stiglitz, J., 1983. Prizes and incentives: towards a general theory of compensation and competition. *Bell J. Econ.* 14, 21–43.
- Schotter, A., Weigelt, K., 1992. Asymmetric tournaments, equal opportunity laws, and affirmative action: some experimental results. *Q. J. Econ.* 107 (2), 511–539.
- Sheremeta, R., Wu, S., 2012. Testing Canonical Tournament Theory: On the Impact of Risk, Social Preferences and Utility Structure. Mimeo.
- Sheremeta, R.M., 2010. Experimental comparison of multi-stage and one-stage contests. *Games Econ. Behav.* 68 (2), 731–747.
- Sheremeta, R.M., 2011. Contest design: an experimental investigation. *Econ. Inq.* 49, 573–590.
- Sheremeta, R.M., Zhang, J., 2010. Can groups solve the problem of over-bidding in contests? *Social Choice Welfare* 35, 175–197.
- Shupp, R., Sheremeta, R.M., Schmidt, D., Walker, J., 2013. Resource allocation contests: experimental evidence. *J. Econ. Psychol.* 39, 257–267.
- Skaperdas, S., Gan, L., 1995. Risk aversion in contests. *Econ. J.* 105, 951–962.
- Sunde, U., 2009. Heterogeneity and performance in tournaments: a test for incentive effects using professional tennis data. *Appl. Econ.* 41 (25), 3199–3208.