RIETI Discussion Paper Series 18-E-066

Competitiveness, Risk Attitudes, and the Gender Gap in Math Achievement

YAGASAKI Masayuki
University of Tokyo

NAKAMURO Makiko
Keio University

Research Institute of Economy, Trade & Industry, IAA
The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Discussion Paper Series 18-E-066
October 2018

Competitiveness, Risk Attitudes, and the Gender Gap in Math Achievement[1]

YAGASAKI Masayuki
University of Tokyo

NAKAMURO Makiko
Keio University

Abstract

In this paper, we investigate how competitiveness and risk attitudes are related to math achievement among middle school students. We conduct an experiment at six public middle schools in Japan to collect incentivized measures of competitiveness and risk attitudes and merge them with an administrative dataset containing information on students' cognitive achievements. The results from the experiment show that girls are less competitive and exhibit greater risk aversion compared to boys, which are in line with the previous literature. We find that competitiveness is positively correlated with math achievement conditional on students' prior achievements and demographics, while greater risk aversion is associated with higher math achievement (but not with reading and English). Taken together, the results indicate that the gender differences in competitiveness are widening the gender gap in math achievement, but that the gender differences in risk attitudes contribute to narrowing it.

Keywords: Gender, Competitiveness, Risk attitudes, Math, Experiment
JEL classification: C91, I20, I24, J16

[1]This research was conducted as a part of the project "Research on the Improvement in Resource Allocation and Productivity among the Healthcare and Education Service Industries" undertaken at the Research Institute of Economy, Trade and Industry (RIETI). The research was financially supported by MEXT KAKENHI (Grant-in-Aid for Scientific Research (A)) Grant Number 16H02022. We are grateful to Hidehiko Ichimura and Yasuyuki Sawada for helpful comments and encouragement during this project. Mitsunosuke Morishita provided us an excellent research assistance at the experiment. We would like to thank Ryuichi Tanaka for detailed comments that have led to the improvement of the paper. We would also like to thank Hideo Akabayashi, Daiji Kawaguchi, Nobuyoshi Kikuchi, Masao Ogaki, Fumio Ohtake, Hiroko Okudaira, and seminar participants at RIETI, Keio University, Chiba University, Kansai Labor Economics Workshop, and JEA 2018 Autumn Meeting (Gakushuin University) for their thoughtful comments and the ensuing discussions. All remaining errors are attributable to the authors.

1 Introduction

Over the last century, there have been substantial improvements in the educational outcomes of female students and they are now attaining higher education at rates similar to or higher than male students in many developed countries (Goldin, Katz, and Kuziemko, 2006). Despite this, we still observe girls performing worse than boys on standardized math examinations. For example, the 2015 Program for International Student Assessment (PISA) finds that boys outperform girls in math by 8 score points on average across OECD countries; Boy's advantage at the mean is statistically significant in 28 countries and economies that participated in PISA (OECD, 2017).

This gender gap in math achievement has gotten particular attention in economics for at least two reasons. First of all, in contrast to other subjects such as reading, math skills and preparation serve as a good predictor of future labor market outcomes. For example, Joensen and Nielsen (2009, 2016) exploit an institutional reduction in the costs of acquiring advanced high school math in Denmark and provide evidence that choice of a more math-intensive high school specialization has a causal effect on future labor market earnings. Secondly, it is also thought that mathematical proficiency does not just benefit individuals but also is considered crucial to drive economic growth and create innovation (Hanushek and Kimko, 2000; Jamison, Jamison, and Hanushek, 2007). Gaining a better understanding of how the gender math gap arises is important to come up with potential policies that improve girls' math performance and narrow the gap.

The objective of this paper is to investigate how gender-linked behavioral traits such as competitiveness and risk attitudes are related to math achievement among middle school students. In doing so, we examine to what extent the gender gap in math is attributable to gender differences in competitiveness and risk attitudes. There is a broad consensus in the experimental literature that women, on average, are less competitive (e.g., Niederle and Vesterlund, 2007) and exhibit greater risk aversion (e.g., Eckel and Grossman, 2002) as compared to men. These two noncognitive behavioral traits may be important in the production of cognitive achievements. As Heckman (2006) argues, noncognitive traits could cause people to endogenously create environments during childhood that foster faster cognitive development. As for competitiveness, for example, students who are more competitive may compete for grades with their peers and improve their cognitive achievements through rivalry. In particular, it is suggested in the previous literature that competitiveness depends on tasks and that the gender differences in competitiveness is clearly observed when the assigned task is male stereotypical like math (see Niederle and Vesterlund, 2011 for a survey). Indeed, Niederle and Vesterlund (2010) is the first to suggest the possibility that the gender differences in competitiveness in part explain the gender math gap. They argue that math may be seen as a competitive discipline because answers are either 'right' or 'wrong' and because math is highly predictive of future earnings. From these, it can be inferred that the gender differences in competitiveness (in male stereotypical tasks) could be an important factor that explains the gender gap in math achievement.

As for risk attitudes, for example, students who are more willing to take risk may attempt to attain better grades with minimal efforts by betting on what problems will be on the exams, which may be inefficient learning strategy in the long run for their skill formation. Or more traditionally, as first suggested by the theoretical work of Lehvari and Weiss (1974), if the returns to educational investment in the future are uncertain, students who are risk averse may lower educational investment which results in lower achievement.

In spite of these various potential mechanisms under which risk attitudes affect the production of cognitive achievements, an empirical relationship between risk attitudes and cognitive achievements is relatively unexplored. If risk attitudes are related to math achievement, it is potentially the case that the gender differences in risk attitudes are related to the gender gap in math achievement.

To this end, we conduct an incentivized experiment at six public middle schools in Japan to collect measures of competitiveness and risk attitudes and merge them with an administrative dataset containing information on students' cognitive achievements. We find that, as predicted, competitiveness is positively correlated with math achievement conditional on students' prior achievements and demographics, while greater risk aversion is associated with higher math achievement (but not with reading and English). Since girls are less competitive and exhibit greater risk aversion compared to boys, the results indicate that the gender differences in competitiveness are widening the gender gap in math achievement, but that the gender differences in risk attitudes contribute to narrowing it.

1.1 Related Literature

First of all, our paper is related to the empirical literature of the production of cognitive achievements (e.g., Todd and Wolpin, 2003, 2007; Cunha and Heckman, 2008). In particular, Cunha and Heckman (2008) construct a dynamic structural model in which cognitive and noncognitive skills evolve jointly and estimate its production function parameters. Even though our approach in this paper is not structural, the paper examines how noncognitive behavioral traits such as competitiveness and risk attitudes are related to the production of cognitive achievements. To the best of our knowledge, this is the first paper which tackles such a question.

Second, the paper adds to the growing literature of behavioral economics of education (e.g., Koch, Nafziger, and Nielsen, 2014; Lavecchia, Liu, and Oreopoulos, 2016). Especially, recent literature accumulates mounting evidence showing that competitiveness is predictive of educational outcomes outside the lab. Buser, Niederle, and Oosterbeek (2014) investigate whether competitiveness explains academic track choice of middle school students in the Netherlands. They find that competitiveness predicts the choice of math-heavy specializations in high school and the gender gap in specializations is largely accounted for (about 20%) by the gender differences in competitiveness. For high school students, Almas et al. (2016) show that competitiveness correlates with choosing the college track in Norway and Buser, Peter, and Wolter (2017) show that competitiveness can explain a significant portion of the gender gap in math-intensive specialization choices in Switzerland. Similarly, Zhang (2013) provides evidence that students who are more inclined to compete are more likely to take a competitive entrance exam for high school in China. Aside from educational choices, recent evidence suggests that competitiveness is predictive of other outcomes such as earnings and investment behavior. [1] In contrast to this literature, our focus is on cognitive achievements, especially math, rather than the educational choices such as academic track choice. As Niederle and Vesterlund (2010) first hypothesized, we will see that competitiveness is positively associated with math achievement (but not with reading and English), explaining part of the gender gap in math.

Starting from a theoretical work by Lehvari and Weiss (1974), the relationship between risk attitudes and educational outcomes is a long-standing area of active research in economics. Traditional view is that risk aversion is inversely associated with educational outcomes since uncertainty in returns to education depresses educational investment (e.g., Belzil and Leonardi, 2007, Checchi, Fiorio, and Leonardi, 2014). Recent literature in experimental economics complements this view. In Buser, Niederle, and Oosterbeek (2014), the authors find that risk attitudes itself is predictive of academic track choices. They report that the more risk averse students are less likely to choose more math-heavy specializations in high school and about 16% of the gender gap in track choices can be explained by the gender differences in risk attitudes. Tannenbaum (2012) analyzes a data sample from the Fall 2001 math SAT and finds that women skip significantly more questions than men. He attributes this difference primarily to gender differences in risk aversion since, in SAT, students are penalized for incorrect answers but not for leaving questions blank. He argues that the gender gap in questions skipped can explain up to 40% of the gender gap in SAT scores. Similarly, using an experiment, Baldiga (2013) finds that women answered significantly fewer questions than men when the wrong answer was penalized, but not when there was no penalty. [2] In contrast to the literature which supports the view that risk aversion is negatively related to educational outcomes, we show that risk aversion is positively related to math achievements.

Finally, the paper adds to the literature on the gender gap in math. A wide range of theories has been proposed to explain the gender gap in math. These theories can be classified into two

[1] Reuben, Sapienza, and Zingales (2015) link the starting salary and industry choice of MBA students and find that competitive individuals earn 9 % more than their less competitive peers. Furthermore, they find that gender differences in tournament entry account for about 10 % of the gender gap in earnings. See also Reuben, Wiswall, and Zafar (2015), Buser, Geijtenbeek, and Plug (2015). Berge et al. (2015) show that competitiveness predicts investment choices of entrepreneurs in Tanzania.

[2] There is also a recent stream of experimental literature that investigate the relationship between risk attitudes and innate cognitive ability (e.g., Frederick, 2005; Burks et al. 2009; Dohmen et al. 2010; Benjamin, Brown, and Shapiro 2013). These studies suggest that risk aversion is negatively related to cognitive ability. However, Andersson et al. (2016) show that this relationship may be spurious. In their study, they show that by changing the way how risk elicitation tasks are presented, they are able to generate both negative and positive correlations between risk aversion and cognitive ability. They argue that cognitive ability is related to behavior error rather than to risk preferences.

broad categories: biological theories such as innate differences in spatial ability, brain development, and theories arguing the importance of societal factors such as differential treatment by parents and teachers, stereotypical threat etc (see Halpern et al., 2007 for a survey). Obviously, sorting out the relative importance of biological versus societal explanations is important since these two imply different policy implications. However, the objective of this paper is not to contribute to that discussion. Rather than that, our objective of this paper is to address the validity of the argument that encourages women to "lean-in" (Sandberg, 2013): women should be more competitive and take on more risk. [3] Our results suggest that, at least from the viewpoint of the gender gap in math achievement, encouraging girls to become more risk tolerant may backfire and result in the loss of an advantage of girls in the production of math achievement.

The remainder of the paper unfolds as follows. Section 2 describes the data collection and experimental procedures. Section 3 presents benchmark analysis. We first demonstrate that there is a significant gender gap in math conditional on prior achievements. We then see that there are significant gender differences in experimental measures such as competitiveness, confidence, and risk attitudes and asses to what extent gender differences in competitiveness can be attributed to gender differences in confidence and risk attitudes. In Section 4, we examine whether competitiveness and risk attitudes correlate with math achievement. We also provide subsample analysis by gender and the regression results of reading and English achievements. Section 5 concludes.

2 Background and Data Collection

We invited 8th-grade students of all 6 public middle schools within the same city of Saitama prefecture, a large part of the Kanto metropolitan area in Japan. Schools are geographically located within 12 km radius. Approximately two months prior to the experiment (Feb 2 through 13, 2017), the authors directly visited all schools and explained the schedule, setting, and financial incentive of the experiment in detail. Students were distributed a letter about details of the experiment to families and a parental consent form, and were required to return a signed consent form by about two weeks.[4]

After all, we received 848 students' parental consent forms (out of a possible 1080) and finally 811 students (389 male, 422 female) from 30 classes participated in our experiment, which were accounted for 75% out of the entire 8th-grade students.[5] To prevent the detailed information on

[3]We are inspired by the discussion of Shurchkov and Eckel (2018) on this part. A related question is whether women should "lean-in" to negotiate more (e.g., Exley, Niederle, and Vesterlund. 2016).

[4]However, the students were not informed on the specific task of the experiment at that time to prevent students from self-selecting into the participation in experiments, based on their favorite tasks. The parental consent form included the same information given to the students. Teachers, except for the principle, were not fully informed about the experiments to make sure students did not find out about the purpose of this experiment.

[5]According to the official statistics, the total numbers of 8th-grade students at the beginning of 2016 academic semester was 1108. However, we excluded 28 students from this calculation who (i) students who were absent on the day of the standardized exam; (ii) students who transferred from/to other schools after the day of the standardized

the experiments from spreading to other schools, we set up the experiments and collected data within three consecutive days, March 21, 22, and 23, 2017.

Students who participated in the experiment received, on average, 1,022JPY (=10USD), with minimum of 500JPY (=5USD) and maximum of 3,400JPY (=34USD), including a fixed participation fee, 500JPY (=5USD). It should be noted that, due to administrative and educational reasons, students were paid by the combination of bookstore gift cards and regular gift cards (called "QUO card" which can be used in many stores, such as convenience stores, drugstores, restaurants, and gas stations, etc). Although students were informed that they were paid with gift cards in advance, they left uninformed of how much they were paid with bookstore gift cards or how much regular gift cards. Since either gift cards can be easily cashed at a cash voucher shop or anywhere, it is unlikely that paying in gift cards, not cash, will cause a potential problem for our results. These gift cards were mailed to each student three months after the experiments, although it was later than the initial schedule (one month after the experiment) due to the unexpected accident on the postage.

2.1 Experiment

Each day on March 21, 22 and 23, 2017, the experiment was conducted after school and it took about an hour. Students were randomly assigned to 44 classrooms in 6 schools, ranging in size from 11 to 28 of them each. To prevent copying the answers from neighbors, students were asked to sit in every other seat in the classroom. We, with assistance of two Research Assistants (RAs) per classroom, administered the experiment for about 60 minutes, including the short survey. To see how experimental environments affect individual decision makings, we used a between-subjects 2 $\times$ 2 design and randomly manipulate environments in the classrooms.[6] The environments differed in the visibility of the choices (private vs public), and the experimental peer groups (same-sex vs mixed gender), as explained below.

The visibility of the choices. We randomly assigned students to choose their choices in the experiment in "public" situations or in "private" situations. In the public treatment, students were announced that their choices during the experiment would be made public to the students who were participating in the experiment in the same room by our research assistants at the end of the experiment. In the private treatment, the choices would be kept private throughout the experiment as in the standard literature.

The experimental peer groups. We randomly assigned students to participate in the experiment with same-sex peer groups or mixed-sex peer groups. This treatment concerns the gender composition in the room where the experiments take place. Students were randomly assigned a room either with same-sex peers or with mixed-sex peers.

exam; and (iii) students who belonged to special education classrooms.

[6] We stratified students by school and gender.

These treatments are designed to see how social image concerns as well as the presence of opposite sex peers affect economic decision making among middle school students which is conceptually similar motivation with Bursztyn, Fujiwara, and Parrais (2017), Buser, Ranehill, and Veldhuizen (2017), and Yagasaki (2018). Eventually, however, we see no statistically significant impacts across any treatments. [7]This suggests that our experimental measures such as competitiveness and risk attitudes are robust to these treatments, enabling us to pool the samples in the following analysis.

The experiment basically follows the standard design of Niederle and Vesterlund (2007). The experiment consisted of five rounds, one of which was randomly selected for payment. In the first three rounds, participants were asked to solve as many as possible mazes in three minutes. Mazes was chosen as a task because an addition task (Niederle and Vesterlund, 2007), a natural task for investigating the link between competitiveness and math achievement, was unavailable for administrative reasons. Thus, we choose mazes since it is stereotypically male tasks as math is. The experiment was conducted using paper and pencil. An example of a maze is shown in Figure 1.

The incentive structure of each round is laid out below.

Round 1: Piece Rate. Students would receive 50 points for each maze correctly solved.

Round 2: Compulsory Tournament. Students were randomly divided into groups of three, and a student who solved the maze most among the three can obtain 150 points per each but the remaining two could not get any points at all. Students were not informed about who they were assigned into the same group as themselves throughout the experiment. If the number of mazes solved were tied at the first place, the winner were chosen randomly.

Round 3: Choice. Students were asked to choose either piece rate or tournament before performing task. If they were to choose piece rate, they would get 50 points per maze solved correctly. If they were to choose tournament, they would get 150 points per maze solved correctly if there score exceeded that of remaining two of the group members in round 2, otherwise they would receive no payment. In case of ties the winner were chosen randomly.

Round 4: Submitting Piece Rate Performance. No maze task was performed here. Students were asked to choose either piece rate or tournament to apply their round 1 piece rate performance. If they were to choose piece rate, they would receive the same payment as they did in round 1. If they were to choose tournament, they would get 150 points per maze if there round 1 score exceeded that of remaining two of the group members in round 2, otherwise they would receive no payment. In case of ties the winner were chosen randomly.

[7]Detailed analysis of this part is under preparation and available upon request.

Round 5: Lottery. Students were asked to pick one option among a sure payoff of 400 points and five 50/50 lotteries: 500 or 350, 600 or 300, 700 or 250, 800 or 200 and 900 or 100 (points). (See Table 1.) For lotteries 1-5, the expected payoff increases linearly with risk, as represented by the standard deviation. Note that lottery 6 has the same expected payoff as lottery 5 but with a higher standard deviation. These lotteries are designed so that higher number of the choice of a lottery implies greater preference for risks.[8] The outcome of the lottery was determined by flipping a coin at the end of the experiment if this round was randomly chosen for compensation.

In rounds 3, 4 and 5, students in the public treatment were announced that the choice of that round would be made public to the peers in the same room, if it was randomly chosen for compensation, at the end of the experiment. Finally, students answered a detailed questionnaire including questions on confidence, psychological attributes and demographics such as family patterns, parental employment status and the number of siblings. Confidence measures were elicited by asking students to guess their relative rank in round 1 and round 2 performances of their group of three. If their guess were correct, they receive 100 points for each. [9]

2.2 Administrative Data

A few months after the experiment, we obtained several administrative data from the local government and matched with the data collected through the experiment. Firstly, we are allowed to access standardized test scores that the local government of Saitama prefecture administered every academic year. This standardized test, started from 2015, was constructed as panel data, tracking down the same students over time. Therefore, one of the greatest advantages of accessing this dataset is we are able to employ a value-added specification of the education production function. The value-added specification adds measures of cognitive achievements in previous academic year as controls. The aim is to account for unobservables that potentially bias the estimates. Including prior year achievements provides sufficient statistics for all historical school and family inputs and students' genetic endowments etc. [10] Secondly, even though we are able to control for the historical input measures by using the information on prior cognitive achievements, we still need the information on demographics and the current state of inputs. We address this issue by the following three ways. One, as mentioned, there are some information on demographics such as

[8] The last column in Table 1 represents implied CRRA range corresponding to each chosen lotteries. The intervals are determined by assuming $u(x) = x^{1-r}$ and calculating the value of r that would make the individual indifferent between the lottery s/he chose and the two adjacent lotteries. Theoretically, individuals with $r > 0$ can be classified as risk averse, $r < 0$ as risk loving and $r = 0$ as risk neutral.

[9] The questionnaire also asks questions on empirical norms. For instance, it asks each student 'what fraction of boys/ girls in your school who participated into the experiment do you think choose tournament in round 2'. If the guess is correct, then the student gets 100 points. These questions are designed to elicit students' belief about their gender stereotype.

[10] See Todd and Wolpin (2003, 2007) for detailed discussions.

family patterns, parental employment status and the number of siblings in the questionnaire collected during the experiment. Two, in administering the standardized tests, students are requested to answer a series of questionnaires, including students' information on age in months, and cram school attendance etc. Three, we access the administrative data that the local government of the city collected every year, such as whether students' guardians receives public assistances and the subsidy for school lunch and school supplies, both of which are the proxy of family wealth. Finally, the important feature of the standardized test is employing the Item Response Theory (IRT) in estimating students' cognitive skills more precisely (for details, see Embretson and Reise, 2000). Contrary to the Classical Test Theory (CCT), the IRT is successful to separate the difficulty level of problems on the test from the difference in students' cognitive skills. In addition, skill estimates of IRT at different times are mapped in a common scale so that the IRT scores of the same student are comparable across different time periods. An important drawback of IRT, however, is that if a student gets either zero or perfect test score, a skill estimate of IRT diverges to negative or positive infinity. Consequently, for these two cases, IRT fails to yield a skill estimate and the data is coded as some symbol to indicate what has happened. In order to address this censoring issue, we mainly use Tobit model in the following analysis.

3 Descriptive Analysis

In this section, we describe basic characteristics of the students who participated in the experiment. Descriptive statistics of variables we use in our main analysis are displayed by gender in Table 2. To keep the sample constant, we had to drop 67 students because at least one of these key variables are missing for those students. This leaves us with a sample of 744 students (345 boys, 399 girls).

3.1 The Gender Gap in Math Achievement

Even though our primary focus is on math test score, it is useful to see test scores on reading and English by gender as well. It is widely known that girls traditionally exceeds boys in overall middle school performance. Indeed, as displayed in Table 2, in both 8-th grade and 9-th grade, girls are outperforming boys in reading and English. As suggested by Goldin, Katz, and Kuziemko (2006), this may be due to the later maturation of boys. However, if that is true, it is mysterious why we see girls not performing better than boys in math as reported in Table 2.

In order to understand the gender effects on achievements more precisely, we estimate regression models. Table 3 reports the results of Tobit regressions using 9-th grade IRT scores of each subject as dependent variables. As described above, we use Tobit model to account for the censoring issue due to the use of IRT.[11] Again, columns (1) to (3) show that girls are on average better at reading

[11] Most of the results are not sensitive to the normality assumption imposed on Tobit model. As a robustness check, we also implement other estimation methods such as OLS by dropping censored data and censored LAD

and English compared to boys, whereas boys are on average better at math than girls, even though the estimated difference in math is not statistically significant. Columns (4) to (6) additionally include cognitive achievements in the previous year.[12] The estimated coefficient of the female dummy in column (4) is negative and significant at the 1% level, implying that boys are likely to achieve greater improvement in math than girls. Given the growth in math of a typical student is about .5 as displayed in Table 2, the coefficient on female dummy variable (.128) amounts to roughly 25% of one year growth in math of a typical student. The statistical significance of female coefficients of reading and English disappear after controlling prior achievements as shown in column (5) and (6). In addition to prior achievements, Column (7) to (9) adds students' demographic variables, which deemed to affect student achievements and are often controlled in standard education production functions (see Todd and Wolpin, 2003, 2007). One of the variables that represent family wealth is a dummy variable that takes one if a students' parents receives either public assistances, or the subsidy for school lunch and school supplies, zero otherwise. Moreover, we also control for family patterns and parental employment status. Family patterns are classified into three type of household; (i) nuclear family (i.e. a couple and a child(ren)), (ii) single parent and a child(ren) and (iii) other. Parental employment status is expressed as a series of dummy variables that correspond to information on who are engaged on a job in household (father, mother, both, or other). An important control, which could be regarded as a current input measure, is the dummy variable taking one if a student attends for-profit private cram schools outside of formal education, making very important role for students to prepare for the entrance examination of high schools, zero otherwise. Further, widely known determinants of student achievements, the number of siblings and student's age in month are also included in our estimation. According to columns (7) to (9), the coefficients are quantitatively and qualitatively similar even after controlling for these variables. In the following empirical analysis, we investigate what factors lie behind the gender gap in math achievement observed in column (7) in Table 3.

3.2 The Gender Differences in Experimental Measures

Table 2 reports average choices and performance in the experiment by gender. Consistent to most of the literature, we find that boys are significantly more likely to enter the tournament than girls are. In our sample, boys are approximately 18 percentage points more likely to choose tournament in round 3. For round 4, boys are 9 percentage points more likely to choose tournament. We also find that boys are on average better in performing mazes in both round 1 and round 2 and the differences are statistically significant.

estimator developed in Powell (1984). The results do not change qualitatively. These results are available upon request.

[12]This value-added specifications include lagged test scores not only in math but also in reading and English. If students allocate their resources, such as time and concentration, to maximize the overall cognitive achievements, not a test performance for a particular subject, it is more convincing to control for the prior own achievement outcomes in reading and English as well as math.

Confidence and risk attitudes also follow the previous literature and we find that boys are more confident about their relative performance in both round 1 and round 2 and more risk-seeking than girls. For confidence, in Table 2 we see that both boys and girls guess their relative performance to be higher when it comes to round 2. This may reflect the effect of learning between rounds 1 and 2. For risk attitudes, Table 2 shows that boys choose a more risky lottery than girls on average than girls do and the difference is statistically significant.

In summary, our sample exhibit the standard patterns of gender differences that we observe in most of the literature. However, it is not clear to what extent the gender differences in tournament entry is attributable to the gender differences in performance, confidence and risk attitudes in our sample. Therefore, we move on to the regression analysis in the next analysis.

Table 4 reports the results of OLS regression of tournament entry in round 3. All specifications include school fixed effects and treatment fixed effects. Column (1) shows that girls are 18 percentage points less likely to enter the tournament than boys, when only controlling for school and treatment fixed effects. Column (2) shows that adding performance in round 2, the difference in performance between rounds 1 and 2, and 8th-grade cognitive achievements reduce the gender effect by 3.1 percentage points (compare columns (1) and (2)). The reduction is as expected, given the gender differences in the number of mazes correctly solved in our sample. Notably, among 8th-grade cognitive achievements, math is the only subject which is (marginally) significant.

In column (3), we add the guessed ranks of rounds 1 and 2 as measures of confidence. We see that adding confidence measures causes the gender effect to drop slightly from 14.9 to 13.4 percentage points. On the other hand, we see a substantial drop in the gender effect when we add the choice of lottery which is a measure of risk attitudes. Comparing columns (3) and (4), adding the lottery choice in round 5 reduces the coefficient of female dummy by 5.2 percentage points (from 13.4 to 8.2 percentage points). Finally in column (5) and (6) we include the dummy of round 4 choice of tournament entry, hereinafter called "submitting the PR", to control other possible factors that influence tournament entry such as feedback aversion. Although submitting the PR significantly predicts tournament entry in round 3, we see almost no effect on the gender effect. Column (6) adds individual controls. Individual controls include dummies of low socioeconomic status, dummies of cram school attendance, dummies of family patterns, dummies of parental employment status, age in months and the number of siblings. Controlling all variables leaves 8.1 percentage points gender gap in tournament entry which is statistically significant at a 5% level.

Overall, the middle school students in our sample exhibit significant gender differences in competitiveness but the point estimate is relatively small (about 8 percentage points) after controlling performance, confidence and risk attitudes. In particular, we see measures of confidence do not have a large impact on the gender gap in tournament entry, whereas the risk attitudes do eliminate substantial portion of the gender effect. This is in contrast with the literature such as Niederle and Vesterlund (2007) and Buser, Niederle, and Oosterbeek (2014) in which authors conclude that significant amount of the gender differences in tournament entry is driven by the gender difference in confidence, whereas the risk attitudes do not have a large impact on the gender differences in tournament entry once controlling for confidence.[13] On the other hand, the results are in line with, for example, Gillen, Snowberg, and Yariv (2016) in which authors argue that differences in risk aversion, rather than confidence, account for the gender gap in their study.

4 Main Results

This section reports our main results of this paper. In Table 5, we estimate different specifications of Tobit model with 9th-grade math achievement as the dependent variable. All specifications include 8th-grade cognitive achievements (which are reported in the table), performance in round 2 of the experiment, the difference in performance between rounds 1 and 2, guessed ranks of rounds 1 and 2, submitting the PR, school fixed effects, treatment fixed effects, and the same set of individual controls as in the previous analysis (which are not reported in the table).

Column (1) shows that girls' math achievement are on average significantly lower than that of boys conditional on prior cognitive achievements.[14] In columns (2) to (4), we add measures of competitiveness and risk attitudes. Note that the female coefficient remains negative and significant even after controlling additional experimental measures that are considered to be correlated with gender. The magnitude of the coefficients, however, varies across specifications, clarifying which factor brings about the gender gap in math achievement.

We first examine the effect of competitiveness on math achievement. Column (2) reports the result when we add the tournament entry dummy. The coefficient on the tournament entry dummy is positive and statistically significant at 10% level, implying that the tournament entry is positively correlated with math achievement. The magnitude of the coefficient is substantial for two reasons. First of all, since the average growth in IRT math score between 8-th and 9-th grade academic year is about .5 (see Table 2), the coefficient on the tournament entry dummy (.105) amounts to roughly 20% of one-year growth in math of a typical student. Secondly, comparing the coefficients on the female and the tournament entry dummies, we see that the magnitude of the tournament entry is larger than that of being female. Moreover, to emphasize the significance, we add all the controls from column (2) but exclude the female dummy. The resulting coefficient of the tournament entry dummy is .117 (p = .030). Compared to the magnitude of the female dummy in column (1), which is .096 (p = .036), tournament entry dummy predicts 9-th grade math achievement slightly better than the female dummy.[15] In column (8), we further add risk attitudes as controls. The coefficients on the tournament entry dummy increases and statistically significant at 5% level. This means that the tournament entry has an effect independent of risk attitudes and

[13] See Niederle and Vesterlund (2011) for a survey on this line.

[14] The results of column (1) of Table 7 are the same as those of column (7) of Table 5, except we now control for performance in round 2 of the experiment, the difference in performance between rounds 1 and 2, guessed ranks of rounds 1 and 2, submitting the PR, and treatment fixed effects.

[15] The way we discuss here follows Buser, Niederle, and Oosterbeek (2014).

other confounding factors. Therefore, we interpret that competitiveness being positively associated with math achievement. In columns (5) and (6), we confirm that this is also true for both boys and girls, i.e., controlling for risk attitudes and other confounding factors, the coefficient on the tournament entry dummy is positive and significant in the subsamples of boys and girls.

We next assess the impacts of including the tournament entry dummy on the gender gap in math achievement. We bootstrapped changes in the gender coefficient upon adding the tournament entry dummy as in Buser, Niederle, and Oosterbeek (2014). The results are reported in Panel A of Table 6. Pairwise comparisons between columns (1) and (2), and (3) and (4) show that the reductions in the female coefficient by including the tournament entry dummy are statistically significant at 5% level. The results show that the gender differences in competitiveness account for 9.2% to 13.5% of the gender gap in math achievement. In other words, the gender differences in competitiveness is widening the gender gap in math achievement.

Secondly, we examine the effect of risk attitudes on math achievement. Column (3) adds risk attitudes to column (1). The coefficient on the risk attitudes is statistically significant at 10% level. The estimated coefficient on the risk attitudes in column (3) should be subject to omitted variable bias. Namely, more risk tolerant students tend to enter tournament in round 3, and the tournament entry is positively correlated with math achievement, so the coefficient on the risk attitudes is positively biased. Columns (4) controls for the entry dummy. We see that the coefficient on risk attitudes increases, in absolute sense, and becomes statically significant at 5% level. Therefore, the results suggest that greater risk aversion is associated with higher math achievement. The magnitudes are also substantial. The estimated coefficients suggest that a student who chose lottery 1 (the most risk averse) gets .130 to .165 higher math IRT score than one who chose lottery 6 (the most risk tolerant) but is otherwise identical. Again, recalling that the average growth in IRT math score is about .5, these effects are substantial. Columns (5) and (6) show that the coefficients on risk attitudes are statistically significant and quantitatively similar across gender.

The impacts of adding risk attitudes on the female coefficient are reported in Panel B of Table 6. Since girls are more likely to be risk averse and greater risk aversion is associated with higher math achievements, controlling risk attitudes increases, in absolute sense, the female coefficient by 16.7% to 31.3%. The magnitudes are statistically significant at 5% level for both specifications. This means that the gender differences in risk attitudes contribute to narrowing the gender gap in math achievements.

Finally, we investigate the relative impacts of controlling competitiveness and risk attitudes. The bottom row (Panel C) reports the reduction in the female coefficient upon controlling competitiveness and risk attitudes simultaneously. Although the impact is not statistically significant, the difference is negative, suggesting that the impact of controlling risk attitudes is slightly stronger than controlling competitiveness in our sample.

4.1 Effects on Reading and English

Table 7 reports the results of regressions using reading and English IRT scores as dependent variables. Interestingly, the estimated coefficients on the tournament entry dummy and the risk attitudes in these regressions are insignificant except for the coefficient of tournament entry on reading achievement in girls sample. Overall, the results for reading and English achievements are unstable across gender and inconclusive.

5 Conclusion

Despite its importance of math skills in life, recent data indicates that boys continue to outperform girls in many developed countries. Consequently, understanding the sources of the gender gap in math achievement has received particular attention in economics. The aim of this paper is to investigate whether gender-linked behavioral traits such as competitiveness and risk attitudes are predictive of math achievement among middle school students. We find that competitiveness is positively correlated with mathematics achievement conditional on students' prior achievements and demographics, while greater risk aversion is associated with higher math achievement. Therefore, the results indicate that the gender differences in competitiveness are widening the gender gap in math achievement, but that the gender differences in risk attitudes contribute to narrowing it.

These findings have a significant policy implication. One of the most important results in our paper is that greater risk aversion, which is considered as a part of female gender norms, is positively correlated with the production of math achievement. This result is in contrast with the previous findings that greater risk aversion is associated with negative economic outcomes such as STEM career choices (Buser, Niederle, and Oosterbeek, 2014) and math SAT scores (Tannenbaum, 2012). As also pointed out by Blau and Kahn (2017), the result suggests that the gender differences in behavioral traits do not necessarily benefit boys and thus the policies intended to encourage girls to "lean in" (Sandberg, 2013) - girls should be more competitive, and take on more risks, etc. - may backfire and lead to the loss of some of the economic advantages of girls. Therefore, rather than the policies that are designed to change girls' behavioral traits, we think that institutional redesign and behavioral interventions that take the gender differences in behavioral traits as given constitute a potentially more promising approach to close the gender gap in economic outcomes.

References

[1] Almas, Ingvild, Alexander W. Cappelen, Kjell G. Salvanes, Erik . Srensen, and Bertil Tungodden, 2016. "What Explains the Gender Gap in College Track Dropout? Experimental and Administrative Evidence." American Economic Review, 106 (5), 296-302.

[2] Andersson, Ola, Hakan J. Holm, Jean-Robert Tyran, and Erik Wengstrom, 2015. "Risk Aversion Relates to Cognitive Ability: Preferences or Noise?" Journal of the European Economic Association 14, no. 5 : 1129-1154.

[3] Baldiga, K.A., 2013. "Gender Differences in Willingness to Guess." Management Science. 60(2): 434-448.

[4] Benabou, Roland, Jean Tirole, 2002. "Self-confidence and Personal Motivation." Quarterly Journal of Economics 117 (3), 871-916.

[5] Benjamin, Daniel J., Sebastian A. Brown, and Jesse M. Shapiro, 2013. "Who is Behavioral? Cognitive Ability and Anomalous Preferences." Journal of the European Economic Association, 11, 1231-1255.

[6] Belzil, C., and Leonardi, M., 2007. "Can Risk Aversion Explain Schooling Attainments? Evidence from Italy." Labour Economics, 14(6), 957-970.

[7] Berge, Lars Ivar Oppedal, Kjetil Bjorvatn, Armando Jose Garcia Pires, and Bertil Tungodden, 2015. "Competitive in the Lab, Successful in the Field?" Journal of Economic Behavior and Organization,118, 303-317.

[8] Blau, D. Francine, and Lawrence M. Kahn, 2017. " The Gender Wage Gap: Extent, Trends, and Explanations," Journal of Economic Literature, 55 (3), 789-865.

[9] Burks, Stephen V., Jeffrey P. Carpenter, Lorenz Goette, and Aldo Rustichini, 2009. "Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment." Proceedings of the National Academy of Sciences, 106, 7745-7750.

[10] Bursztyn, Leonardo, Thomas Fujiwara, and Amanda Pallais, 2017. "'Acting Wife': Marriage Market Incentives and Labor Market Investments," American Economic Review, 107 (11), 3288-3319.

[11] Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek, 2014. "Gender, Competitiveness and Career Choices." Quarterly Journal of Economics 129 (3):1409-1447.

[12] Buser, Thomas, Noemi Peter, and Stefan Wolter, 2017. "Gender, Competitiveness, and Study Choices in High School: Evidence from Switzerland." American Economic Review: Papers and Proceedings 107 (5):125-130.

[13] Buser, Thomas, Eva Ranehill, and Roel van Veldhuizen, 2017. "Gender Differences in Willingness to Compete: The Role of Public Observability" available at SSRN

[14] Checchi, D., Fiorio, C. V., and Leonardi, M., 2014. "Parents' Risk Aversion and Children's Educational Attainment." Labour Economics, 30(0), 164-175.

[15] Cunha, F., Heckman, J.J., 2008. "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation." Journal of Human Resources. 43 (4), 738-782.

[16] Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde, 2010. "Are Risk Aversion and Impatience Related to Cognitive Ability?" American Economic Review, 100(3), 1238-1260.

[17] Eckel, Catherine C., Grossman, Philip J., 2002. "Sex Differences and Statistical Stereotyping in Attitudes Toward Financial Risk." Evolution and Human Behavior 23(4): 281-295.

[18] Embretson S. E., and Reise, S. P., 2000. "Item Response Theory for Psychologists." Mahwah, NJ : Psychology Press.

[19] Exley, Christine L., Muriel Niederle, and Lise Vesterlund, 2016. "Knowing When to Ask: The Cost of Leaning-in." HBS Working Paper 16-115.

[20] Frederick, Shane, 2005. "Cognitive Reflection and Decision Making." Journal of Economic Perspectives, 19(4), 25-42.

[21] Gillen, B., E. Snowberg and L. Yariv, 2018 "Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study." working paper.

[22] Goldin C., Katz LF, Kuziemko I., 2006. "The Homecoming of American College Women: the Reversal of the College Gender Gap." Journal of Economic Perspectives. 20(4):133-56.

[23] Halpern, D., C. Benbow, D. Geary, R. Gur, J. Hyde and M. Gernsbacher, 2007. "The Science of Sex Differences in Science and Mathematics." Psychological Science in the Public Interest, vol. 8(1), pp. 1-51.

[24] Hanushek, E. and D. Kimko, 2000. "Schooling, Labor-Force Quality and the Growth of Nations. " American Economic Review 90(5):1184-1208.

[25] Heckman, James, 2006. "Skill Formation and the Economics of Investing in Disadvantaged Children." Science, 312(5782): 1900-02.

[26] Jamison, E., D. Jamison, and E. Hanushek, 2007. "The Effects of Education Quality on Income Growth and Mortality Decline. " Economics of Education Review, 26(6): 771-88.

[27] Joensen, J. S. and H. S. Nielsen, 2009. "Is There a Causal Effect of High School Math on Labor Market Outcomes?" Journal of Human Resources 44 (1):171-198.

[28] Joensen, J. S. and H. S. Nielsen, 2016. "Mathematics and Gender: Heterogeneity in Causes and Consequences." Economic Journal, 126:1129-1163.

[29] Koch, A., Nafziger, J., and Nielsen, H.S., 2014. "Behavioral Economics of Education." Journal of Economic Behavior and Organization, 115, 3-17.

[30] Lavecchia, A., H. Liu, and P. Oreopoulos, 2016. "Behavioral Economics of Education: Progress and Possibilities." in Handbook of the Economics of Education, ed. by S. M. Eric A. Hanushek, and L. Woessmann, vol. 5, chap. 1, pp. 1-74. Elsevier.

[31] Lehvari, D., Weiss, Y., 1974. "The Effect of Risk on the Investment in Human Capital. " American Economic Review. 64, 950-963.

[32] Niederle, M. and L. Vesterlund, 2007. "Do Women Shy Away from Competition? Do Men Compete too Much?" Quarterly Journal of Economics 122 (3): 1067-1101.

[33] Niederle, M and L. Vesterlund, 2010. "Explaining the Gender Gap in Math Test Scores: The Role of Competition." Journal of Economic Perspectives 24(2):129-144.

[34] Niederle, M. L. Vesterlund, 2011 "Gender and Competition" Annual Review of Economics, 3, 601-630.

[35] Powell, James, 1984. "Least Absolute Deviations Estimator for the Censored Regression Model." Journal of Econometrics 25, 303-325.

[36] Reuben, Ernesto, Paola Sapienza, and Luigi Zingales, 2015. "Taste for Competition and the Gender Gap among Young Business Professionals." Working paper.

[37] Sandberg, Sheryl, 2013. "Lean in: Women, Work, and the Will to Lead." New York: Alfred A. Knopf.

[38] Shurchkov, O., and C. C. Eckel., 2018. "Gender Differences in Behavioral Traits and Labor Market Outcomes." In S. L. Averett, L. M. Argys, and S. D. Hoffman (eds.), Oxford Handbook on Women and the Economy. Oxford: Oxford University Press.

[39] Tannenbaum, Daniel I., 2012. "Do Gender Differences in Risk Aversion Explain the Gender Gap in SAT Scores? Uncovering Risk Attitudes and the Test Score Gap." working paper

[40] Todd, P., Wolpin, K., 2003. "On the Specification and Estimation of the Production Function for Cognitive Achievement." Economic Journal.113, 3-33.

[41] Todd, P., and Wolpin, K., 2007. "The production of cognitive achievement in children: Home, school, and racial test score gaps." Journal of Human capital 1.1. 91-136.

[42] Yagasaki, Masayuki, 2018. "How Can We Induce More Women to Competitions? Social Image and Prosocial Incentives." in progress.

[43] Yeager, D. S., Henderson, M. D., Paunesku, D., Walton, G. M., D'Mello, S., Spitzer, B. J., and Duckworth, A. L., 2014. "Boring but important: A self-transcendent purpose for learning fosters academic self-regulation." Journal of Personality and Social Psychology, 107(4), 559-580.

[44] Zhang, J., 2013. "Can Experimental Economics Explain Competitive Behavior Outside the Lab?" mimeo, University of California, Berkeley.

A Tables

Table 1: Lotteries in Round 5

Choice(50/50 lottery)	High	Low	Mean	SD	Implied CRRA range
Lottery 1	400	400	400	0	$3.94 < r$
Lottery 2	500	350	425	75	$1.32 < r < 3.94$
Lottery 3	600	300	450	150	$0.81 < r < 1.32$
Lottery 4	700	250	475	225	$0.57 < r < 0.81$
Lottery 5	800	200	500	300	$0 < r < 0.57$
Lottery 6	900	100	500	400	$r < 0$

Notes. This table shows the choice of lotteries in round 5. In the experiments, we show only points associated to each lotteries to the students. We do not show means and standard deviations displayed in the table. The last column represents implied CRRA range corresponding to each chosen lotteries. The intervals are determined by assuming $u(x) = x^{1-r}$ and calculating the value of r that would make the individual indifferent between the lottery s/he chose and the two adjacent lotteries. Theoretically, individuals with $r > 0$ can be classified as risk averse, $r < 0$ as risk loving and $r = 0$ as risk neutral.

Table 2: Descriptive Statistics

	Boys			Girls			Difference
	N	Mean	SD	N	Mean	SD	t-test
IRT scores							
9-th grade math	342	1.40	1.06	394	1.30	1.05	
9-th grade reading	345	1.49	1.33	399	1.82	1.24	***
9-th grade English	343	1.01	1.15	396	1.34	1.13	***
8-th grade math	345	0.92	1.07	399	0.84	0.95	
8-th grade reading	345	0.84	1.07	399	1.19	1.02	***
8-th grade English	345	0.16	1.04	399	0.49	0.99	***
Growth in math	342	0.49	0.62	394	0.48	0.64	
Growth in reading	345	0.66	0.93	399	0.63	0.87	
Growth in English	343	0.86	0.61	396	0.86	0.63	
Experimental variables							
Performance (Piece-rate)	345	6.09	1.85	399	5.51	1.78	***
Performance (Tournament)	345	8.23	2.31	399	7.45	2.31	***
Tournament entry (round 3)	345	0.41	0.49	399	0.23	0.42	***
Tournament entry (round 4)	345	0.24	0.43	399	0.15	0.36	***
Guessed rank (round 1)	345	2.06	0.64	399	2.21	0.59	***
Guessed rank (round 2)	345	1.69	0.71	399	1.78	0.73	***
Lottery	345	4.08	1.80	399	3.04	1.61	***
Parental employment status							
Only father is employed	345	0.16	0.37	399	0.16	0.37	
Only mother is employed	345	0.08	0.26	399	0.05	0.21	*
Both	345	0.76	0.43	399	0.78	0.42	
Other	345	0.01	0.09	399	0.02	0.12	
Family patterns							
Nuclear family	345	0.74	0.44	399	0.78	0.42	
Single parent and a child(ren)	345	0.11	0.31	399	0.06	0.24	*
Other	345	0.15	0.36	399	0.16	0.37	
Other controls							
Cram school attendance	345	0.75	0.43	399	0.72	0.45	
Number of siblings	345	1.24	0.82	399	1.29	0.89	
Age in months	345	173.49	3.35	399	173.34	3.43	
Low SES	345	0.16	0.36	399	0.15	0.36	

Notes. The table reports means and standard deviations of variables by gender based on 744 students. The last column reports gender differences in means where the significance levels are from t-test ; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Gender and Cognitive Achievements

	(1) Math	(2) Reading	(3) English	(4) Math	(5) Reading	(6) English	(7) Math	(8) Reading	(9) English
Female	-0.097	0.310***	0.319***	-0.128***	0.081	0.032	-0.119**	0.087	0.031
	(0.082)	(0.094)	(0.086)	(0.047)	(0.058)	(0.047)	(0.047)	(0.058)	(0.047)
8th-grade math				0.697***	0.360***	0.157***	0.683***	0.356***	0.154***
				(0.035)	(0.042)	(0.034)	(0.035)	(0.042)	(0.034)
8th-grade reading				0.050	0.419***	0.173***	0.054*	0.420***	0.174***
				(0.031)	(0.038)	(0.030)	(0.031)	(0.038)	(0.031)
8th-grade English				0.243***	0.382***	0.750***	0.230***	0.375***	0.751***
				(0.035)	(0.042)	(0.034)	(0.035)	(0.043)	(0.034)
School FE	√	√	√	√	√	√	√	√	√
Individual controls							√	√	√
Observations	744	744	744	744	744	744	744	744	744

Notes. Coefficients are from Tobit regressions using 9-th grade achievements as dependent variables. All specifications include school fixed effects. Individual controls include dummies of low socioeconomic status, dummies of cram school attendance, dummies of family patterns, dummies of parental employment status, age in months and the number of siblings. Robust standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Determinants of Tournament Entry

	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.180***	-0.149***	-0.134***	-0.082**	-0.084**	-0.081**
	(0.034)	(0.036)	(0.036)	(0.036)	(0.035)	(0.036)
T-PR		0.003	0.010	0.011	0.018	0.017
		(0.014)	(0.015)	(0.015)	(0.014)	(0.014)
Tournament		0.037***	0.014	0.013	0.007	0.007
		(0.010)	(0.011)	(0.010)	(0.011)	(0.011)
8th-grade math		0.054**	0.041	0.039	0.040	0.041*
		(0.026)	(0.025)	(0.025)	(0.025)	(0.025)
8th-grade reading		0.027	0.021	0.025	0.028	0.027
		(0.023)	(0.022)	(0.021)	(0.022)	(0.021)
8th-grade English		-0.012	-0.013	-0.008	-0.007	-0.010
		(0.025)	(0.025)	(0.025)	(0.025)	(0.025)
Guessed rank R1			-0.051	-0.050	-0.019	-0.013
			(0.032)	(0.031)	(0.032)	(0.032)
Guessed rank R2			-0.123***	-0.092***	-0.094***	-0.099***
			(0.028)	(0.028)	(0.027)	(0.028)
Lottery				0.059***	0.053***	0.052***
				(0.009)	(0.010)	(0.010)
Submitting the PR					0.148***	0.151***
					(0.052)	(0.052)
School and treatment FE	√	√	√	√	√	√
Individual controls						√
Observations	744	744	744	744	744	744

Notes. Dependent variable: tournament entry dummy of round 3. The table presents coefficients from OLS regressions. All regressions control for school fixed effects and treatment fixed effects. Tournament is performance in the round 2 compulsory tournament. T-PR is the difference in performance between the round 2 tournament and the round 1 piece rates. Submitting the PR is the tournament entry dummy of round 4. Individual controls are dummies of low socioeconomic status, dummies of cram school attendance, dummies of family patterns, dummies of parental employment status, age in months and the number of siblings. Robust standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Math Achievement and Behavioral Traits

	(1) Full	(2) Full	(3) Full	(4) Full	(5) Boys	(6) Girls
Female	-0.096**	-0.083*	-0.120**	-0.109**		
	(0.046)	(0.047)	(0.049)	(0.049)		
Entry		0.105*		0.130**	0.120*	0.157*
		(0.055)		(0.056)	(0.071)	(0.090)
Lottery			-0.026*	-0.033**	-0.038**	-0.037*
			(0.014)	(0.014)	(0.019)	(0.022)
8th-grade math	0.667***	0.662***	0.669***	0.663***	0.597***	0.727***
	(0.036)	(0.037)	(0.036)	(0.037)	(0.051)	(0.052)
8th-grade reading	0.050	0.048	0.049	0.046	0.079*	0.017
	(0.032)	(0.032)	(0.032)	(0.032)	(0.047)	(0.043)
8th-grade English	0.225***	0.226***	0.223***	0.225***	0.244***	0.203***
	(0.034)	(0.034)	(0.034)	(0.033)	(0.046)	(0.047)
Observations	744	744	744	744	345	399
Pseudo R^2	0.401	0.403	0.403	0.405	0.456	0.382

Notes. Coefficients are from Tobit regressions using 9-th grade math achievement as a dependent variable. All specifications include performance in round 2 of the experiment, the difference in performance between rounds 1 and 2, submitting the PR, guessed ranks of rounds 1 and 2, school fixed effects, treatment fixed effects, and individual controls. Individual controls include dummies of low socioeconomic status, dummies of cram school attendance, dummies of family patterns, dummies of parental employment status, age in months and the number of siblings. Robust standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Bootstrap Results of the Reductions in the Female Coefficient

	Columns	Difference	Percentage change	p-value
Panel A: Competitiveness	(1)-(2)	0.013	13.5%	0.031
	(3)-(4)	0.011	9.2%	0.023
Panel B: Risk attitudes	(1)-(3)	-0.016	-16.7%	0.027
	(2)-(4)	-0.026	-31.3%	0.011
Panel C: Competitiveness + Risk attitudes	(1)-(4)	-0.013	-13.5%	0.189

Notes. This table reports the results of bootstrap for the reduction in the female coefficient upon controlling for competitiveness and risk attitudes with 10,000 repetitions. p-value is equal to the number of repetitions divided by 10,000 in which the reduction points toward the opposite direction.

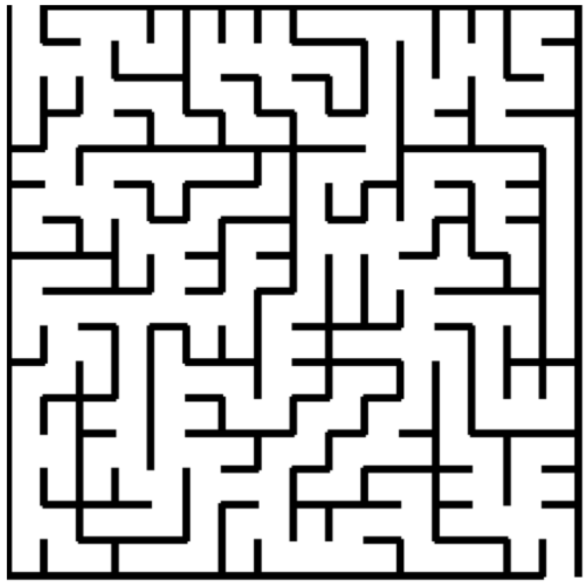
Table 7: Reading, English Achievements and Behavioral Traits

	(1)	(2)	(3)	(4)	(5)	(6)
		Reading			English	
	Full	Boys	Girls	Full	Boys	Girls
Female	0.085			0.046		
	(0.058)			(0.049)		
Entry	0.043	-0.044	0.156*	0.026	0.055	0.016
	(0.063)	(0.091)	(0.089)	(0.055)	(0.074)	(0.086)
Lottery	-0.020	-0.034	-0.000	0.001	-0.017	0.014
	(0.016)	(0.023)	(0.026)	(0.014)	(0.019)	(0.022)
8th-grade math	0.343***	0.278***	0.380***	0.145***	0.062	0.233***
	(0.044)	(0.063)	(0.059)	(0.035)	(0.051)	(0.048)
8th-grade reading	0.414***	0.395***	0.439***	0.171***	0.216***	0.142***
	(0.040)	(0.062)	(0.051)	(0.032)	(0.046)	(0.044)
8th-grade English	0.378***	0.460***	0.313***	0.752***	0.768***	0.729***
	(0.048)	(0.069)	(0.061)	(0.037)	(0.052)	(0.050)
Observations	744	345	399	744	345	399
Pseudo R^2	0.335	0.338	0.349	0.433	0.456	0.426

Notes. Coefficients are from Tobit regressions using 9-th grade achievements as dependent variables. All specifications include performance in round 2 of the experiment, the difference in performance between rounds 1 and 2, submitting the PR, guessed ranks of rounds 1 and 2, school fixed effects, treatment fixed effects, and individual controls. Robust standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

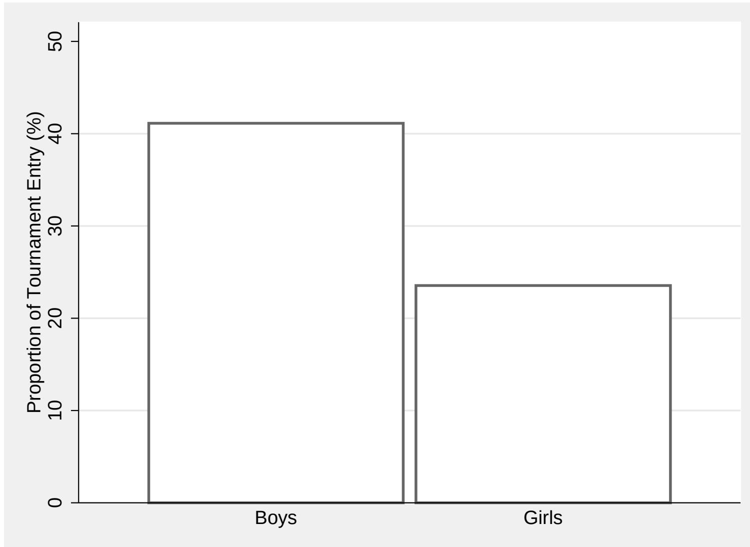
B Figures

Figure 1: Example of a maze



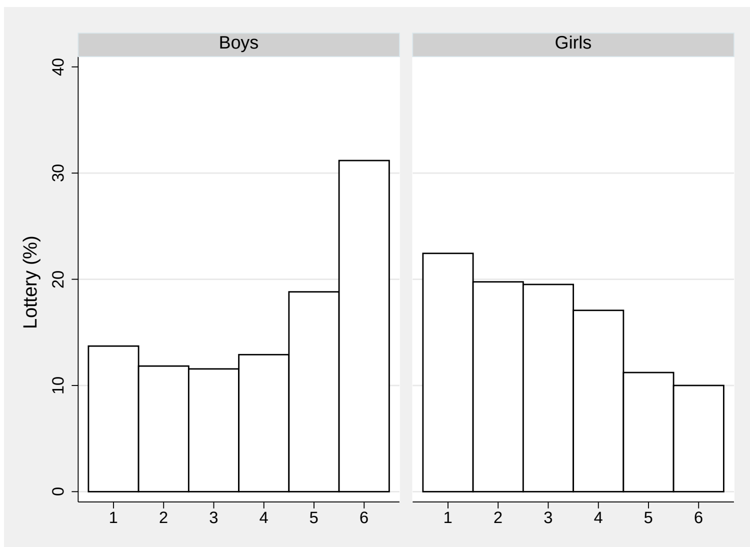
Notes. This is an example of a maze used in the experiment.

Figure 2: Tournament Entry in Round 3



Notes.

Figure 3: Lottery Choice in Round 5



Notes.