Gender, competitiveness and career choices

Buser, T.; Niederle, M.; Oosterbeek, H.

Published in:
The Quarterly Journal of Economics

DOI:
10.1093/qje/qju009

GENDER, COMPETITIVENESS, AND CAREER CHOICES*

Thomas Buser
Muriel Niederle
Hessel Oosterbeek

Gender differences in competitiveness have been hypothesized as a potential explanation for gender differences in education and labor market outcomes. We examine the predictive power of a standard laboratory experimental measure of competitiveness for the later important choice of academic track of secondary school students in the Netherlands. Although boys and girls display similar levels of academic ability, boys choose substantially more prestigious academic tracks, where more prestigious tracks are more math- and science-intensive. Our experimental measure shows that boys are also substantially more competitive than girls. We find that competitiveness is strongly positively correlated with choosing more prestigious academic tracks even conditional on academic ability. Most important, we find that the gender difference in competitiveness accounts for a substantial portion (about 20%) of the gender difference in track choice. *JEL* Codes: C9, I20, J24, J16.

I. Introduction

A recently emerging literature documents large gender differences in competitiveness based on laboratory experiments (see Croson and Gneezy 2009; Niederle and Vesterlund 2011). While women shy away from competition, men often compete too much. It has been hypothesized that these gender differences in competitiveness may help explain gender differences in actual education and labor market outcomes. Evidence supporting this hypothesis is, however, thin. Bertrand (2011) attributes this to the rather new research agenda and the difficulty in finding databases that combine a good measure of competitiveness with real outcomes. This article aims to fill this gap. To assess the relevance of competitiveness for education outcomes—and gender differences therein—we link a standard experimental measure of competitiveness with the later important choice of academic

*We thank the staff of the four schools that allowed us to collect the data from their pupils that we use in this article. We thank Lawrence Katz and four anonymous referees for their comments. We also thank Gerrit Bloothooft for classifying students into socioeconomic groups on the basis of their first names. Nadine Ketel and Boris van Leeuwen provided excellent research assistance. Niederle thanks the NSF for financial support. We gratefully acknowledge financial support from the University of Amsterdam through the Speerpunt Behavioural Economics.

The Quarterly Journal of Economics (2014), 1409–1447. doi:10.1093/qje/qju009.
Advance Access publication on May 8, 2014.

track of secondary school students in the Netherlands. The tracks from which these students can choose vary considerably in their prestige and their math and science intensity.

Math and science intensity is one of the most significant dimensions of gender differences in educational choices. In many Organisation for Economic Co-operation and Development countries, girls are less likely than boys to choose math- and science-heavy courses in secondary education.[1] In the United States, a gender difference in math and sciences manifests itself at the college level, where women are significantly less likely than men to graduate with a major in science, technology, engineering, or mathematics.[2] While in U.S. high schools girls take as many advanced math and science classes as boys and perform at similar levels on average (Goldin, Katz, and Kuziemko 2006), girls are still underrepresented among extremely high-achieving high school students in math (Ellison and Swanson 2010). The main reason to be concerned about gender differences in math and sciences is that the choice of math and science classes is a good predictor of college attendance and completion (Goldin, Katz, and Kuziemko 2006). Performance in mathematics also predicts future earnings (for evidence and discussion, see Paglin and Rufolo 1990; Grogger and Eide 1995; Brown and Corcoran 1997; Altonji and Blank 1999; Weinberger 1999, 2001; Murnane et al. 2000; Schroter Joensen and Skyt Nielsen 2009; Bertrand, Goldin, and Katz 2010).

It has been suggested that gender differences in math are explained by gender differences in ability. However, Ellison and Swanson (2010) provide compelling evidence that the gender imbalance among high achieving math students in the United States is not driven by differences in mathematical ability alone. Moreover, among equally gifted students, males are much more likely to choose a math-heavy college major (see LeFevre, Kulak, and Heymans 1992; Weinberger 2005).

1. We will show that in the Netherlands boys are significantly more likely to take math classes in secondary school than are girls. In France, like in the Netherlands, secondary school children decide on which sets of classes to enroll in, and girls are less likely to choose the math- and science-heavy options (see http://www.insee.fr/fr/themes/tableau.asp?ref_id=eduop709&reg_id=19). The same is true for Denmark (Schroter Joensen and Skyt Nielsen 2011), Switzerland (http://www.ibe.uzh.ch/publikationen/SGH2003_d.pdf), and Germany (Roeder and Gruehn 1997).

2. See http://nces.ed.gov/pubs2009/2009161.pdf.

Intuitively, it seems plausible that competitiveness could be a relevant trait to explain entry into fields such as sciences and mathematics, which are male dominated and viewed as competitive.[3] There is, for example, evidence of a low tolerance for competition among women who drop out of math-intensive college majors and engineering (see Felder et al. 1995; Goodman Research Group 2002). This evidence is fairly casual, however, and may suffer from reverse causality. Women who drop out of science and engineering may search for explanations such as the negative aspect of a competitive environment.

In the Netherlands, secondary school students at the preuniversity level choose at the end of the third year of their six-year program (at age 15) between four study tracks: a science track, a health track, a social sciences track, and a humanities track. These four tracks are clearly ranked in terms of math intensity and academic prestige as follows (in descending order): science, health, social sciences, and humanities. This choice of academic track strongly correlates with the choice of major in tertiary education later in life.

We use the measure of competitiveness introduced by Niederle and Vesterlund (2007). Participants perform a real effort task first under a noncompetitive piece rate incentive scheme and then a competitive tournament scheme. In the tournament, a participant competes against three group members, and only the subject with the highest performance receives a payment. Participants do not receive any information about the performance of others, including whether they won the tournament. For their final task, participants choose between the competitive tournament scheme and the noncompetitive piece rate scheme. This choice serves as a measure of the participants' willingness to compete. The typical result is that controlling for performance,

3. Experiments have shown that for women both the performance in as well as the willingness to enter competitive environments is reduced when the competition group includes males (Gneezy, Niederle, and Rustichini 2003; Balafoutas and Sutter, 2012; Niederle, Segal, and Vesterlund 2013). Similarly, Huguet and Regner (2007) show that girls underperform in mixed-sex groups (but not in all female groups) in a test they were led to believe measures mathematical ability. Furthermore, some studies find that the strong gender differences in competitiveness found in mathematical tasks are sometimes but not always attenuated when assessed in verbal tasks (e.g., Kamas and Preston 2010; Dreber, von Essen, and Ranehill 2014; Wozniak, Harbaugh, and Mayr 2014; see Niederle and Vesterlund 2011 for an overview).

males are more likely to enter the tournament than females. Niederle and Vesterlund (2007) show that about one third of the gender gap in tournament entry can be accounted for by gender differences in confidence. Risk attitudes are often shown to only play a minor role. Since both confidence and risk aversion may not only account for tournament entry but also influence students' academic track choice, we administered incentivized measures of students' confidence and risk attitudes. Because we are concerned with the choice of prestigious science tracks typically favored by males, we measure the competitiveness of students in a stereotypical male task, arithmetics, which is also the task used in Niederle and Vesterlund (2007) and most of the resulting literature. This measure of competitiveness has proven to be robust across different settings and subject pools (Niederle and Vesterlund 2011).

We administered our experiment in four schools in and around Amsterdam, measuring students' competitiveness a few months before they chose their study track. To avoid problems of reverse causality, it is important to measure competitiveness before students have different and potentially influential experiences resulting from their choices. After the school year, the schools provided us with the track choices of their students as well as with their grades. Because grades may not be the most accurate predictor of ability, we asked the students for their own perceptions of their mathematical ability.

The students in our sample exhibit the expected gender differences. Although the academic performance of girls (including math grades) is at least as good as that of boys, boys choose substantially more prestigious academic tracks than girls. Also, while the performance of boys and girls on the experimental task is very similar, boys are twice as likely as girls to choose the competitive payment scheme. Our first finding is that the choice of the tournament scheme in the experiment is significantly positively correlated with the prestige and math and science intensity of the chosen academic track. Being competitive bridges around 20% of the distance between choosing the lowest and the highest ranked track. The effect of competitiveness is comparable in size to the effect of being male. Our main finding is that gender differences in competitiveness can account for 20% of the gender gap in the prestige and math and science intensity of the chosen academic track, controlling for grades and perceived mathematical ability. We show that the effect of competitiveness is not driven by a possible effect of confidence or risk attitudes on academic track choice. When we control for the experimental measures of confidence and risk attitudes, the decision to enter the tournament still closes the gender gap in the prestige of chosen tracks by a significant 16%. These results not only demonstrate the external relevance of the concept of competitiveness but also validate the specific measure of competitiveness provided by Niederle and Vesterlund (2007).

The remainder of the article proceeds as follows. The next section describes the collection of the data and the variables. Section III provides more details of the academic tracks from which the students have to choose. We then present the results from the study in three stages. First, in Section IV we document significant gender differences in the prestige (and math intensity) of the track choices made by the students in our sample and show that they cannot be explained by gender differences in ability. Second, in Section V we document significant gender differences in competitiveness and assess the extent to which these differences can be attributed to gender differences in confidence and risk attitudes. Finally, in Section VI we examine whether competitiveness correlates with track choice and assess to what extent gender differences in competitiveness can account for gender differences in the prestige and math and science intensity of chosen tracks. Section VII discusses alternative interpretations of the tournament entry measure, and Section VIII concludes.

II. Study Design

II.A. Data Collection

We invited secondary schools in and around Amsterdam to participate in a research project investigating the determinants of study track choices of students in the pre-university level of secondary school. We demanded one class hour (45 or 50 minutes) of all grade 9 classes at the pre-university level. The invitation letter stated that students would participate in an in-class experiment and be paid depending on their choices. For detailed instructions, see the Online Appendix. We describe the Dutch school system and the study track choice in Section III.

Four schools gave us access to their students, one in the city of Amsterdam and three in cities close to Amsterdam. In each school, we captured all students in grade 9 of the pre-university level for a total of 397 students in 16 classes. Because the schools are geographically dispersed, we do not worry that students received information about the experiment from students in other schools. For any given school, experiments in different classes were administered on the same day, often at the same time. The data collection in the schools took place in March, April, and May 2011. The participants were paid a week after the experiment through sealed envelopes. They earned an average of €5.55, with a minimum of zero and a maximum of €25. There was no fixed participation fee.

After the end of the school year, the schools provided us with the students' final grades in grade 9 and their track choices for the last three years of high school. For 35 students we do not have such a track choice. For 20 of these students, we can use information about their expected track choice obtained through the short questionnaire at the end of the experiment.[4] We drop the remaining 15 students for whom we have neither a definite choice nor a clear answer from the questionnaire. We have to drop an additional four students from the analysis because they showed up late to class and missed part of the experiment, two students because their questionnaires were incomplete and they therefore lack key control variables, and 14 students because we did not obtain their grades. This leaves us with a sample of 362 subjects.

Although a sample of four schools can hardly be representative for the total of over 500 schools in the Netherlands that offer the pre-university level, the four schools appear to be average on several dimensions. First, as we will discuss in more detail later, the track choices by gender in the four schools are close to the national averages. Second, with 51%, the proportion of girls in our sample is close to the national average of 53% at the pre-university level. Third, the numbers of students in the four schools are (around) 700, 800, 1,500, and 2,000. The average secondary school size in the Netherlands is close to 1,500 (CBS 2012, p. 80). Fourth, the average grades on the nationwide exams in the final year in the four schools are 6.2, 6.2, 6.3, and 6.8, where the national average for the pre-university level equals 6.2. Finally, the pass rates on the final exam in the four schools are 0.87, 0.87,

4. For the students for whom we have both the definite track choice and the intention stated in the questionnaire, the questionnaire answer accurately predicts the final choice in 93% of the cases.

0.91, and 0.95, where the national average for the pre-university level equals 0.88.

II.B. *Experimental Variables and Student Characteristics*

1. Competitiveness. Our measure of competitiveness is taken from Niederle and Vesterlund (2007). Participants perform a real task in three rounds, first under a noncompetitive piece rate scheme and then under a competitive tournament scheme. Participants then choose which of the two payment schemes to apply to their third and final performance. This choice serves as our measure of their competitiveness.

The task of the experiment is to add up sets of four two-digit numbers for three minutes. The performance in each round corresponds to the number of correctly solved problems. In each round participants received envelopes that contained a sheet of 26 problems. There were always three versions of the 26 addition problems to prevent copying from neighbors. After having read out the instructions that were on top of the envelopes and answering questions (if any), the experimenter gave the signal that subjects could open the envelopes and start the addition problems. Participants were not allowed to use calculators but could use scratch paper. Once the three minutes of solving problems were over, subjects had to drop the pen and stand up.

Participants were informed at the start of the experiment that they would perform in three rounds, one of which would be randomly chosen for payment at the end of the experiment through the roll of a die in front of the classroom. Participants received detailed instructions on each round only immediately before performing in the task in that round. Participants did not receive any information about their own performance or the performance of others at any point during the experiment. Only a week later, when participants were paid, could they make inferences about their relative performance.

In round 1, participants were paid for their performance according to a non-competitive piece rate of 25 euro-cents per correctly solved problem. In round 2, they performed in a tournament against three competitors. The competitors were randomly selected by computer among students from the same class after the end of the experiment. The person with the largest number of correctly solved problems received €1 per correct problem and the others received no payment. In case of a tie, the winner was randomly determined.

In round 3, participants could choose which of the two payment schemes would be applied to their performance. Like in round 1, a participant who chose the piece rate received 25 cents per correct problem. A participant who selected the tournament would win if her new round 3 performance exceeded the round 2 performance of her three competitors. In case of a tie, the winner was randomly determined. Therefore, just like in Niederle and Vesterlund (2007), the choice of payment scheme was an individual decision as a subject could not affect the payoffs of any other participant.[5]

2. Confidence. The decision whether to enter the tournament in round 3 may depend on the students' beliefs about their relative performance in their group of four competitors. We therefore elicited those beliefs. Specifically, we asked the students to guess their rank in the round 2 tournament, from 1 (best) to 4 (worst) of their group of four. If their guess was correct, they received €1.[6]

3. Risk attitudes. The decision to enter the tournament in round 3 may also depend on the students' risk attitudes. We elicited risk attitudes using two separate measures from the experimental literature. First, following Eckel and Grossman (2002), subjects picked one option among a sure payoff of €2 and four 50/50 lotteries with increasing riskiness and expected payoffs: 3 or 1.5; 4 or 1; 5 or 0.5; 6 or 0. The outcome of the lottery was determined by the roll of a die at the end of the experiment. Second, we asked subjects "How do you see yourself: Are you generally a person who is fully prepared to take risks or do you

5. There are several advantages to having participants compete in round 3 against the previous round 2 tournament performance. First, the performance of a subject who chose the tournament is evaluated against the performance of other subjects in a tournament. Second, the choice of compensation scheme of a subject should not depend on the choices of other players. Third, the participant causes no externality to another subject. Hence motives such as altruism or fear of interfering with someone else's payoff play no role.

6. When two subjects have the same number of correctly solved additions, they receive the same rank. For example, if two subjects are tied for first place, they are both ranked first and receive €1 if their guessed rank is equal to 1. The next best subject is ranked third.

try to avoid taking risks?". The answer is on a scale from 0 ("unwilling to take risks") to 10 ("fully prepared to take risk"). Dohmen et al. (2011), using representative survey data from Germany, find that this simple nonincentivized risk question predicts both incentivized choices in a lottery task as well as risk taking across a number of contexts, including holding stocks, being self-employed, participating in sports, and smoking. Lonnqvist et al. (2010) find the question to be much more stable over time than lottery measures of risk attitudes.

4. Subjective ability. We collected two subjective ability measures in the postexperimental questionnaire. First, we asked the students to rank themselves on mathematical talent compared to other students in their year (and school) on a scale from 1 (the best 25%) to 4 (the worst 25%).[7] Second, we asked the students how difficult they find it to pass their math class on a scale from 0 (very easy) to 10 (very hard). Although these questions may yield a better assessment of mathematical ability than grades, they could in addition be a measure of confidence, which in turn could influence study track choices. Indeed, it has been found that conditional on academic performance, boys are more confident in their relative ability than girls (Eccles 1998), a difference that seems greatest among gifted children (Preckel et al. 2008).

5. Student characteristics. For each student, we obtained their first name, gender, birth date, and expected track choice. Though we did not collect any socioeconomic background data on the students in our sample, we have their names. Bloothooft and Onland (2011) show that in the Netherlands, first names are strongly predictive of social class, income, and lifestyle, and

7. This was phrased as three yes/no questions: "Do you think your mathematics ability is in the top 25% of your year?", ". . . top 50% of your year?", ". . . top 75% of your year?" A student who answered all three questions with no was automatically assumed to be in the bottom 25%. We had 44 students who answered no to all questions. A student who answers yes to one of the questions also should answer yes to the next (if one is in the top 25%, one is also in the top 50%). Sixty-seven students, however, switched back to no. For these students, we count the first yes as their true answer. Clearly "wrong" answers consist of the yes/yes/no and yes/no/yes patterns. All other patterns can be rationalized by (i) students truly understanding the question or (ii) misreading the question and answering yes only to their own quartile. There are 10 answers that follow the yes/yes/no pattern and 0 that follow the yes/no/yes pattern.

they develop a classification of names into 14 categories. We applied this classification to our sample (Table A.I in the Online Appendix lists the 14 categories and their proportions). The schools provided us with their grades and track choices, which we merged by name and birth date. In the Dutch school system, grades are expressed on a scale from 1 (worst) to 10 (best), where 6 is the first passing grade. We use the students' grades at the end of ninth grade to construct three objective ability measures. The first is grade point average (GPA), calculated as the average of all grades. The second is their grade for mathematics. Grades need not be a perfect predictor of mathematical ability which is why we also collected subjective ability measures. The third is each student's relative math grade compared to the rest of her class. The rank of each student is equal to 1 plus the number of students with a strictly better grade. To compute this rank, we include all students in our sample for whom we have grades, including the students we had to drop for the final results. We then normalize this measure by dividing by the number of students in each class.

In all our analyses, we standardize all nonbinary control variables to have mean zero and a standard deviation of 1. This facilitates the comparison of the magnitudes of the coefficients of different control variables. Table A.I in the Online Appendix provides the mean and standard deviations of all our control variables.

III. The Study Track Choice

The students in our study are drawn from the population of Dutch secondary school students who are enrolled at the pre-university level. In the Dutch school system, tracking first takes place when students go from primary school (grades 1 to 6) to secondary school, normally at age 12. There are three levels: around 20% of students graduate from the six-year pre-university level, 25% from the five-year general level, and 55% from the four-year vocational level. Who enrolls at which level is to a large extent determined by the score on a nationwide achievement test administered at the end of primary school. Girls are somewhat more likely to go to the pre-university level, making up 53% of the students (CBS 2012). In the first three years at the pre-university level, students are taught in the same class of around 25 students for all subjects during the entire school year.

Although the composition of classes may change from year to year, this is in none of the four schools in our study based on further ability tracking. Different subjects are typically taught by different teachers.

Halfway through the six years of secondary school, at the end of grade 9, students at the pre-university level have to choose one of four study tracks:

- the science-oriented track Nature & Technology (NT)
- the health-oriented track Nature & Health (NH)
- the social sciences–oriented track Economics & Society (ES)
- the humanities-oriented track Culture & Society (CS)

Each student can select any track, though low grades in specific subjects may lead to teachers recommending other tracks.

Table I shows the subjects offered in each study track and the number of teaching hours assigned to each subject during the last three years of secondary school, grades 10–12. Mathematics is the only subject taught at a different level in each track, whereby D is the most advanced version followed by B, A, and C. The order of math and science difficulty is therefore NT > NH > ES > CS. There is a strong correlation between the study track a student picks in secondary school and the choice of major in tertiary education. Most NT graduates go on to study a subject in science and engineering, NH graduates often opt for health-related subjects, ES graduates often choose a major in economics and business or law, and most CS graduates choose a subject in the humanities, social sciences, or law.[8]

We have two sources of information about the study track choices of students in our sample. In the questionnaire, we asked students which track they expected to choose. The schools

8. The tertiary education distribution by study track is as follows: Of students in the NT track, 64% study science and engineering, 15% economics and business, 9% a subject in the humanities, and 7% health care. For NH students, 48% study in health care, 18% in science and engineering, 9% in social sciences, and 8% in economics and business. For ES students, 46% study in economics and business, 20% in law, 19% in social sciences, and 8% in humanities. For CS students 34% study in social sciences, 30% in humanities, and 20% in law. For details see Table A.III in the Online Appendix. Some studies actually restrict entry to certain tracks or courses within tracks. For example, medical schools require NT or NH; to study math, having taken at least Math B in high school is required. Source: http://www.cbs.nl/nl-NL/menu/themas/onderwijs/publicaties/artikelen/archief/2007/2007-2193-wm.htm (Statistics Netherlands); the data are from 2006.

TABLE I

Subjects and Teaching Hours per Academic Track

Nature & Technology: NT	Nature & Health: NH
Mathematics B 600	Mathematics A 520
Physics 480	Biology 480
Chemistry 440	Chemistry 440
Nature, life and technology 440	Nature, life and technology 440
or IT 440	or geography 440
or biology 480	or physics 480
or mathematics D 440	
Economics & Society: ES	**Culture & Society: CS**
Mathematics A 520	Mathematics A or C 480
Economics 480	History 480
History 440	Art 480
Management and organization 440	or philosophy 480
or geography 440	or modern foreign language 480
or social studies 440	or Greek or Latin 600
or modern foreign language 480	Geography 440
	or social studies 440
	or economics 480

Notes. The table lists the subjects per track and the number of teaching hours per subject during the last three years of the pre-university level. In addition all students take the following non-track-specific subjects: Dutch (480 hours), English (400), second foreign language, Latin or Greek (480), social studies (120), general natural sciences (120), culture (160), sports (160). The students spend roughly half their time on track-specific subjects and half on common subjects.

Source. Ministry of Education, Culture and Science of the Netherlands.

provided us with information about their actual choices made several months later. Two of the four schools in our sample allow students to pick combined tracks. Of the 173 students in those two schools, 64 students choose the NT/NH combination and 18 the ES/CS combination. In the NT/NH track, students take Mathematics B but physics is not required. In the ES/CS track, students replace one of the CS electives with the economics course. As such, the combined tracks are somewhat in between the pure tracks, though a little closer to NT and ES, respectively. For the main analysis of this article we use for the students in combined tracks the chosen track as stated in the questionnaire.[9]

9. All of the students who picked ES/CS chose ES or CS in the questionnaire. All of the students who picked NT/NH chose NT or NH in the questionnaire with the exception of one student who chose CS. We treat this student as a CS student when using the stated track to place students that chose a combination track into "pure"

However, since one can argue that the NT/NH track is closer to NT, and the ES/CS track closer to ES, we reestimate all regressions using this alternative definition of track choice in the Online Appendix. As a further robustness check, in the Online Appendix, we show results where we treat NT/NH and ES/CS as separate categories.[10] All our results remain qualitatively the same in both specifications.

Related to the difficulty and amount of mathematics and sciences in the curriculum, NT is generally viewed as the most challenging and prestigious study track, followed by NH and ES. CS is seen as the least demanding and least prestigious study track. In other countries in which students can choose study tracks in school, the prestigiousness of the tracks is also often highly correlated with their math intensity (see for example Pautler 1981 for France). The prestigiousness of study tracks is also related to the likelihood of going to university: students in the NT track have an 81% chance to go to university, followed by NH students (72%) and ES students (69%), whereas only 60% of CS students go to university.[11]

The ordering of study tracks is related to the academic performance of the students in our sample. The top panel of Table II shows mean values of measures for students' objective and subjective ability by study track. According to all five of our ability measures, the students who choose NT score higher than the students who choose NH, who in turn score higher than the students who choose ES. Students who choose CS score lowest on four of the five measures.

The ordering of tracks by prestige is also reflected in the students' opinions. In the questionnaire, we asked the students to rank the four study tracks according to "Which track do the best students pick?" The bottom panel of Table II shows that their responses concur with the general opinion. A majority of over 70% of students believes NT is chosen by the best students. A majority of students ranks the NH study track second and ES third. More

tracks. See Table A.II in the Online Appendix for the number of students who pick NT/NH or ES/CS.

10. For these two analyses we drop an additional 20 students. These are all the students for whom we have not received a final track choice from the schools and used the questionnaire answer instead. The questionnaire, however, did not allow for combination tracks.

11. Source: CBS (2010); the data are from students graduating in 2009.

TABLE II

Student Ability and Perceived Prestigiousness of Academic Tracks Given Their Academic Track Choice

By chosen track	NT	NH	ES	CS	Difference
GPA (1–10)	7.12	7.09	6.65	6.61	.000
Math grades (1–10)	7.25	6.73	6.20	6.21	.000
Math relative (0–1)	0.23	0.35	0.46	0.49	.000
Math difficulty (0–10)	1.95	3.62	4.90	5.30	.000
Math quartile (1(best)–4)	1.52	1.98	2.50	2.67	.000
All: prestige (% rank)	1.48 (71%)	2.13 (57%)	2.64 (60%)	3.67 (81%)	
Boys: prestige (% rank)	1.43 (75%)	2.24 (57%)	2.59 (56%)	3.68 (82%)	
Girls: prestige (% rank)	1.52 (68%)	2.03 (57%)	2.71 (64%)	3.66 (80%)	
Observations	102	89	128	43	

Notes. Top rows: Average characteristics of subjects who chose that track among the 362 students in our sample. Grades are out of 10 with higher numbers being better grades. Math relative is the rank of a student's math grade compared to the other students in his class normalized by the number of students, where 1 refers to the student with the lowest math grade. Math difficulty is the answer to the question "How difficult is it for you to get a passing grade in mathematics?" and goes from 0 (very easy) to 10 (very hard). Math quartile is the answer to a question asking the students to rank themselves on mathematical ability compared to other students in their year (and school) on a scale from 1 (in the best 25%) to 4 (in the worst 25%). These measures of ability were collected before the study track choice took effect. The last column reports *p*-values from Kruskal-Wallis tests for the null hypothesis of equality of distributions across tracks. Bottom rows: average perceived ranking of prestigiousness of academic tracks, and in parentheses, the fraction of students who rank that track first (for NT), second, third, or fourth for NH, ES, and CS, respectively.

than 80% of students rank CS as the track chosen by the weakest students. Boys and girls generally agree on the ranking of the tracks. We also asked students to rank the four study tracks in terms of future earnings. The picture that emerges is very similar.[12]

In the remainder of the article, we order the tracks from most to least prestigious (or most to least math- and science-intensive): NT > NH > ES > CS. As a robustness check, we use in the Online Appendix for each student the ranking they gave to their chosen track in terms of which tracks the best students choose. That is, if a student ranked, say, CS as the track chosen by the best students, followed by ES, NH, and NT (so, the reverse order) and chose track CS for herself, we categorize that student as choosing the most "own prestigious" track (rank 4). If this student chose

12. The exact question was "With which track do you think you would earn most in 10 years' time? Rank the tracks from 1 to 4 where 1 means that you would earn most if you chose that track and 4 that you would earn least if you chose that track." This question was only asked to students in two of the four schools and the percentages are therefore based on 181 observations. Fifty percent think that NT gives the best salary prospects, 27% think NH, 20% ES, and 2% CS.

ES, we would rank her choice as 3 and so on. The main results remain qualitatively the same.

IV. Gender Differences in Prestige of Chosen Tracks

Track choices in the Netherlands differ markedly between the sexes (see the left panel of Table III). Overall, boys are more likely to choose more prestigious study tracks. Compared to girls, boys are almost twice as likely to choose the most prestigious science-oriented track, NT, and only a third as likely to choose the least prestigious humanities-oriented track, CS. The NH track is chosen slightly less often by boys than girls and the ES track is chosen in equal proportion.[13] The fact that girls are disproportionately more likely to choose CS has prompted a debate with the minister for Education even proposing to eliminate that track altogether. This idea was ultimately rejected and the tracks remain as they are.[14]

The right panel of Table III shows track choices by gender in our sample of 362 students. The pattern is similar to the pattern observed in the national statistics. The most prestigious NT track is much more popular among boys than girls, and the opposite holds for NH. The ES track is slightly more popular among boys than girls, and girls are much more likely than boys to choose the least prestigious track, CS. These gender differences in track choice are highly significant ($p = .000$; Fisher's exact test). Note that in our sample, boys and girls are as likely to choose one of the science tracks (NT or NH) compared to one of the society tracks (ES or CS). The girls in our sample are quite similar to the national statistics where 49% of girls pick a science track. However, the national statistics show boys to be slightly more likely to pick a science track than our sample (60% versus 52%).

These gender differences in track choice cannot be explained by academic performance. The first three rows of Table IV show that girls in our sample have a significantly higher GPA than boys, and there is no significant gender difference in the absolute

13. Source: CBS (2012). In the numbers above, we categorize NT/NH-combi students as NT students and ES/CS-combi students as ES students. In Table A.II of the Online Appendix, we compare the choices of students in the Netherlands to the ones in our sample using different ways to allocate combination students.

14. Source: http://nos.nl/artikel/203421-minister-wil-onderwijs-reorganiseren.html and http://nos.nl/artikel/268284-raad-niet-minder-profielen-havovwo.html.

TABLE III

Academic Tracks by Gender: National Statistics and Our Sample (Percentages)

	National statistics		Our sample	
	Boys	Girls	Boys	Girls
NT	43	23	40	17
NH	17	26	12	36
ES	35	32	39	32
CS	5	18	8	15

Notes. In the left panel, we treat NT/NH students as NT and ES/CS students as ES in the national data. In our own sample, of the 22 boys who chose NT/NH, 15 stated NT as their favorite track in the questionnaire, 6 NH, and 1 CS. Of the 42 girls who chose NT/NH, 13 put NT and 29 put NH. Of the six boys who chose ES/CS all six put ES. Of the 12 girls who chose ES/CS, 8 put ES and 4 put CS. We use this information to split them into the four tracks in the right panel.

Source. CBS (2012). The data are from 2012.

TABLE IV

Ability by Gender

	Boys	Girls	*p*-value
GPA (1–10)	6.80	6.97	.008
Math grade (1–10)	6.67	6.59	.491
Math relative (0–1)	0.38	0.37	.885
Math difficulty (0–10)	3.41	4.18	.009
Math quartile (1(best)–4)	1.97	2.25	.032
Number of observations	177	185	

Notes. Average characteristics by gender among the 362 students in our sample. Grades are out of 10 with higher numbers being better grades. Math relative is the rank of a student's math grade compared to the other students in his class normalized by the number of students, where 1 refers to the student with the lowest math grade. Math difficulty is the answer to the question "How difficult is it for you to get a passing grade in mathematics?" and goes from 0 (very easy) to 10 (very hard). Math quartile is the answer to a question asking the students to rank themselves on mathematical ability compared to other students in their year (and school) on a scale from 1 (in the best 25%) to 4 (in the worst 25%). The last column reports *p*-values for the null hypothesis of no differences between boys and girls (we use a *t*-test for equality of means except for math quartile, where we use Fisher's exact test).

or relative grade for mathematics. The last two rows of Table IV show that there are, however, significant gender differences on the two subjective measures of mathematical ability (math quartile and math difficulty), with girls feeling less able than boys.

To more precisely understand the gender differences in the prestige of the chosen study tracks, we estimate ordered probit equations where we order tracks from most to least prestigious: NT > NH > ES > CS. Estimation of an ordered probit model simultaneously provides coefficients of the explanatory variables as well as threshold values separating adjacent tracks. These threshold values divide the standard normal density function into four parts, where the densities of these parts give the predicted probabilities of the respective tracks when all explanatory variables are set equal to 0. Changes in the values of the explanatory variables then move the location of the density function relative to the fixed thresholds. To interpret a coefficient, one has to consider not only the size of the coefficient but also the values of the thresholds. To compare the results from different specifications of the ordered probit regressions, we standardize the coefficient for the female dummy by dividing it by the difference between the estimated ordered probit thresholds of the highest and the lowest tracks. This standardized effect is then the fraction of the distance between lowest and highest tracks that is bridged by switching on the female dummy.

Instead of estimating ordered probit models which exploit the ordering of the four academic tracks, we can also estimate OLS models. In addition to the ordering, applying OLS assumes that the difference in prestige between any two adjacent tracks is the same. This is why coefficients in OLS regressions can be directly compared across regressions and do not have to be normalized by threshold values. Although this is a strong assumption, it turns out that our findings are robust to this modeling choice. In fact the magnitudes of our main results remain unchanged. For our main findings we report results from OLS regressions in the Online Appendix.

The first column of Table V shows that boys are significantly more likely than girls to choose a prestigious track. Being female bridges over 18% of the distance between the most and the least prestigious tracks (this is shown in the penultimate row by $\frac{\text{Female}}{(\text{C3}-\text{C1})}$). Including objective ability variables (column (2)) increases the gender gap to 22% of the distance between the most and the least prestigious tracks. Note that the coefficient on female is larger (in absolute value) than on the GPA. An increase of 1 standard deviation in GPA corresponds to bridging only 13% of the gap between the most and the least prestigious tracks.

When we add students' perceptions about their mathematics ability in column (3), the gender gap shrinks but remains large and highly significant. Although these subjective variables may already be viewed as psychological attributes, it may well be that they produce an additional insight into a students' real mathematical ability compared to grades only. In any case, there is a

TABLE V

Determinants of Academic Track Choice; No Psychological Attributes

	Ordered probit (NT > NH > ES > CS)			NT vs. rest	
	(1)	(2)	(3)	(4)	(5)
Female	−0.325***	−0.443***	−0.319**	−0.222***	−0.187***
	(0.115)	(0.124)	(0.126)	(0.047)	(0.043)
Math grade		0.174	−0.074		0.015
		(0.187)	(0.192)		(0.064)
GPA		0.250**	0.244**		0.024
		(0.098)	(0.097)		(0.031)
Math relative		−0.155	−0.145		−0.050
		(0.152)	(0.152)		(0.053)
Math difficulty			−0.240***		−0.076**
			(0.089)		(0.032)
Math quartile			−0.315***		−0.083***
			(0.074)		(0.025)
Cut 1 (C1)	−1.423***	2.120	−0.625		
Cut 2 (C2)	−0.307**	3.358**	0.714		
Cut 3 (C3)	0.353**	4.113***	1.538		
$\frac{\text{Female}}{(\text{C3}-\text{C1})}$	−0.183***	−0.222***	−0.148***		
N	362	362	362	362	362

Notes. Dependent variable in columns (1) to (3): track choice, where NT > NH > ES > CS. Coefficients are from ordered probit regressions. Dependent variable in columns (4) and (5): dummy variable NT = 1. Coefficients in columns (4) and (5) are from OLS regressions. All regressions control for school fixed effects. Robust standard errors in parentheses. *, ** and *** denote significance at 10%, 5%, and 1%, respectively; p-values for $\frac{\text{Female}}{(\text{C3}-\text{C1})}$ are bootstrapped.

significant gender difference in study track choice, with girls choosing less prestigious tracks than boys.[15]

Table A.IV in the Online Appendix shows that the results are very similar when we classify an NT/NH combined choice as NT and an ES/CS choice as ES, instead of using the students' answer in the questionnaire to attribute combined track choices to one of the four baseline study tracks. The results are also robust to

15. Alternatively, when we use simple OLS regressions, where CS is modeled as a choice of 1 up to NT as a choice of 4, the coefficient on female is –0.277 (std. err. 0.106, $p < .01$) controlling only for school fixed effects. The magnitude of the effect increases to –0.341 (std. err. 0.099, $p < .01$) when we add the controls from column (2) in Table V, which is slightly larger than the coefficient on standardized GPA, which is 0.209 (std. err. 0.071, $p < .01$). When we add all the controls from column (3) the gender coefficient is –0.215 (std. err. 0.094, $p < .05$), again larger than the coefficient on the GPA of 0.193 (std. err. 0.066, $p < .01$).

treating the combined tracks as their own category, where combined tracks are ordered between the baseline study tracks, that is, NT > NT/NH > NH > ES > ES/CS > CS. Using the student-specific ordering and running the same ordered probit specifications, we find that the gender differences are, if anything, slightly exacerbated (see the last three columns of Table A.IV in the Online Appendix).

We also estimate linear probability models of choosing the most prestigious track, NT, compared to any other track, controlling for objective and subjective academic performance. Columns (4) and (5) of Table V shows that girls are around 20 percentage points less likely to choose NT, a significant difference.

V. Gender Differences in Competitiveness

In this section we describe gender differences in competitiveness among the students in our sample and analyze to what extent these are due to gender differences in confidence, risk aversion, and performance. Panels A and B of Table VI report mean values of performance in rounds 1 and 2 and of tournament entry in round 3, separately for boys and girls. In accordance with most of the literature, we find that the boys are significantly more likely to choose the tournament than the girls are. We have 49% of the boys but only 23% of the girls entering the tournament.

This gender gap in tournament entry cannot be explained by gender differences in performance. In round 1, where the students are paid a piece rate, boys perform significantly better than girls. However, in round 2, where payment is based on a tournament, there is no significant gender difference in performance. We then compute for each student the chance to win the tournament in round 2 given their performance and that of their classmates.[16] The average chance of winning the tournament is slightly but not significantly higher for boys than for girls. Provided the performance in round 3 is not lower than in round 2, every student with a chance of winning of 25% or higher has

16. To compute the chance of winning the tournament for each participant, we include all 397 students in our sample, including the 35 students we had to drop for the final results. We use simulations and randomly draw 1,000 different comparison groups of three from a participant's own class. If two performances were tied for first place, a 0.5 win was assigned ($\frac{1}{3}$ in case of three tied performances and 0.25 in case of four).

TABLE VI

DESCRIPTIVE STATISTICS OF TASK PERFORMANCE AND PSYCHOLOGICAL ATTRIBUTES BY GENDER

	Scale	Boys	Girls	p-value
Panel A: performance				
Performance round 1 (piece rate)	number of correct answers	6.60	5.94	.03
Performance round 2 (tournament)	number of correct answers	7.90	7.42	.15
Chance of winning round 2 (tournament)	[0,1]	0.27	0.24	.24
Panel B: competitiveness				
Actual tournament entry	dummy	0.49	0.23	.00
Optimal tournament entry	dummy	0.38	0.35	.59
Panel C: confidence				
Actual guessed rank	1 (best)–4 (worst)	2.14	2.56	.00
Optimal guessed rank	1 (best)–4 (worst)	2.39	2.55	.24
Guesses to be the best	dummy	0.32	0.11	.00
Optimal to guess to be the best	dummy	0.25	0.22	.46
Actual guessed rank is correct	dummy	0.38	0.34	.44
Panel D: risk attitudes				
Lottery choice	1 (no risk)–5 (highest risk)	3.46	2.99	.00
Risk taking	0 (avoid risk)–10 (seek risk)	6.52	5.96	.00
Number of observations		177	185	

Notes. The table reports average values of variables by gender based on 362 students. Panel A: performance is the number of correct answers on addition tasks under piece rate and tournament incentives. Chance of winning round 2 is based on simulations drawing 1,000 different comparison groups of three from a participant's own class. Panel B: actual tournament entry is share choosing the tournament scheme in round 3. Optimal tournament entry is the share that has higher expected payoff under tournament than under piece rate given the round 2 tournament performance. Panel C: actual guessed rank is guessed rank in the round 2 tournament. Optimal guessed rank is the guessed rank in round 2 that maximizes expected payoffs. Guesses to be the best is the share that guessed they were ranked first. Optimal to guess to be the best is the share for whom the guessed rank in round 2 to be the best maximizes expected payoffs. Actual guessed rank is correct is 1 if guessed rank is the correct rank in comparison with three randomly drawn students from one's own class. Panel D: lottery choice is choice between five lotteries increasing in riskiness and expected payoffs. Risk taking is the response to the question whether someone sees her-/himself as someone who is fully prepared to take risks (10, highest) or someone who tries to avoid taking risks (0, lowest). The last column reports p-values from t-tests for continuous variables and from a Fisher's exact test for categorical variables of gender differences.

higher expected earnings when choosing to compete in round 3. This would result in 38% of the boys and 35% of the girls entering the tournament, an insignificant difference. The actual gender gap in tournament entry is significantly larger than this expected gender gap in tournament entry ($p = .01$, Fisher's exact test).[17]

As Panel C of Table VI shows, boys are on average more confident about their relative performance in round 2 of the experiment than are girls. The average guessed rank in their group of four is 2.14 for boys and 2.56 for girls, with the two distributions being significantly different. Moreover, 32% of the boys and 11% of the girls believe that they are the best performers within their group, again a significant difference. To assess the accuracy of these beliefs, we compute for each student the optimal guessed rank, that is, the guess that would have maximized their expected earnings given their own performance and the performances of the other students in their class.[18] There is no significant gender difference in the means of the optimal guessed ranks. There is also no significant gender difference in the share of students for whom it is optimal to guess that they are the best in their group of four competitors.[19]

Boys are also more risk-seeking than girls. Panel D in Table VI shows that boys choose a significantly more risky lottery on average than girls do. On the general risk tolerance question, boys also score significantly higher on average. The correlation between the two risk measures is 0.42 in the whole sample

17. In round 3, subjects who compete solve on average 9.75 correct sums whereas those who do not compete solve 7.92 ($p = .00$). The overall average is 8.57. Neither for the subjects who enter the tournament nor for those who choose the piece rate is performance significantly different between the genders ($p = .25$ and $p = .65$, respectively).

18. We compute the optimal guessed rank through simulation. We randomly draw 1,000 different comparison groups of three from a participants' own class. We include all 397 students in our sample, including the 35 students we had to drop for the final results. We counted the number of times a student ranked first, second, third, and fourth. The mode of the ranks is the best guess as it maximizes expected earnings. If two performances were tied for a place, both guesses were counted as correct.

19. Panel C also shows that 38% of the boys and 34% of the girls guess their rank correctly. These shares are quite similar to the shares of correct guesses in Niederle and Vesterlund (2007), who report that 30% of the men and 38% of the women guess their rank correctly. This indicates that the beliefs of the students in our study about the ability of their competitors are as accurate (or inaccurate) as those of participants in an anonymous laboratory setting.

($p < .01$), and 0.45 and 0.34 in the subsamples of boys and girls, respectively ($p < .01$ in both cases).

To assess to what extent gender differences in tournament entry are due to gender differences in confidence, risk aversion and performance, Table VII reports OLS regression results of tournament entry in round 3. Girls have a 23 percentage point lower probability of entering the tournament than boys, when only controlling for performance in round 1, the difference in performance between rounds 1 and 2, the chance of winning in round 2, school fixed effects, and test version fixed effects (column (1)).

Column (2) shows that adding the guessed rank as a measure of confidence causes the gender effect to drop from 23 to 16 percentage points, which is still a substantial and significant gender difference.[20] Adding the lottery choice variable reduces the gender gap in tournament entry by an additional 3 percentage points to 13 percentage points (compare columns (2) and (3)). Adding the questionnaire-based risk measure reduces the gender gap by 1 further percentage point (compare columns (3) and (4)). Finally, also including measures of objective and subjective academic ability hardly affects the gender gap in tournament entry. Including all the controls leaves a significant gender gap in tournament entry of 12 percentage points (column (5)).

In summary, the secondary school students in our sample exhibit the standard gender difference in competitiveness that has been observed for college students (see Niederle and Vesterlund 2011). Controlling for performance, girls are about 23 percentage points less likely to enter the tournament. Slightly over 30% of this gender gap can be explained by gender differences in confidence. Risk attitudes, whether measured by a lottery choice or a simple questionnaire item, significantly predict tournament entry but reduce the gender gap in competitiveness only by a small amount once we control for confidence.

20. Since the task is a mathematics task, we could alternatively use the students' beliefs about their relative performance in mathematics and their beliefs about their math ability. This, however, reduces the gender gap only by about 5% and a gap of 22 percentage points remains. The coefficient on female is –0.223 (std. err. 0.047, $p < .01$), not very different from the –0.233 from column (1). Adding all measures of beliefs about one's relative performance and math ability does not reduce the coefficient on female compared to just having the belief on tournament performance (guessed rank). Female students are then 15.9 (std. err. 4.4, $p < .01$) percentage points less likely to enter the tournament.

TABLE VII
DETERMINANTS OF TOURNAMENT ENTRY

	(1)	(2)	(3)	(4)	(5)
Female	−0.233***	−0.158***	−0.130***	−0.122***	−0.117***
	(0.047)	(0.045)	(0.045)	(0.044)	(0.045)
Tournament	0.037**	0.011	0.010	0.011	0.006
	(0.015)	(0.014)	(0.014)	(0.014)	(0.014)
T − PR	−0.027***	−0.022**	−0.020**	−0.019*	−0.017*
	(0.011)	(0.010)	(0.010)	(0.010)	(0.010)
Win prob	0.263	0.119	0.102	0.072	0.138
	(0.169)	(0.157)	(0.157)	(0.153)	(0.158)
Guessed rank		−0.205***	−0.200***	−0.182***	−0.169***
		(0.027)	(0.027)	(0.027)	(0.028)
Lottery			0.080***	0.042*	0.040*
			(0.023)	(0.024)	(0.024)
Risk-taking				0.102***	0.107***
				(0.021)	(0.022)
Math grade					0.116*
					(0.065)
GPA					−0.057*
					(0.033)
Math relative					0.020
					(0.051)
Math quartile					0.024
					(0.026)
Math difficulty					0.000
					(0.028)
N	362	362	362	362	362

Notes. Dependent variable: round 3 choice of compensation scheme (1, tournament, and 0, piece rate). The table presents coefficients from OLS regressions. All regressions control for school fixed effects and test version fixed effects. Tournament is performance in the round 2 tournament. T-PR is the difference in performance between the round 2 tournament and the round 1 piece rates. Win prob is the chance of winning the round 2 tournament. Standard errors are in parentheses; *$p < .10$, **$p < .05$, ***$p < .01$.

VI. CAN COMPETITIVENESS ACCOUNT FOR GENDER DIFFERENCES IN PRESTIGE OF CHOSEN TRACKS?

This section reports the main results of the article. We assess whether the experimental variable of competitiveness measured during ninth grade is positively correlated with the prestigiousness and hence math and science intensity of the study track chosen at the end of ninth grade. This study track choice then determines the curriculum for the last three years of secondary school. More important, we assess how much of the gender gap in track choices can be accounted for by gender differences in competitiveness. We then show that these results are robust to the inclusion of controls for confidence and risk aversion. Finally, the results are also robust to the inclusion of controls for socioeconomic background.

We estimate different specifications of ordered probit equations with the ranked academic track as the dependent variable. For different sets of controls variables, we are then interested in how much the coefficient of the female dummy changes when we include a dummy that indicates whether students entered the tournament in round 3 of the experiment. This tells us how much of the gender difference in track choices can be explained by competitiveness (as measured through our experiment). We are also interested in the size and significance of the tournament entry dummy itself. Table VIII reports the results. All specifications include controls for objective and subjective ability (which are reported in the table), and for performance in rounds 1 and 2 of the experiment, the chance of winning the round 2 tournament, school fixed effects, and test version fixed effects (which are not reported in the table).

Columns (1) and (2) report results when we do not add any other controls. The results in column (1) are the same as those of column (3) of Table V, except that we now also control for performance on the experimental task and test version fixed effects. Being female bridges 15% of the gap between choosing the least and most prestigious tracks. We then add the tournament entry dummy in column (2). The coefficient of this dummy is positive and significant at the 5% level. Students who enter the tournament in the experiment are more likely to choose more prestigious academic tracks. The effect size is substantial. In column (2), where we include both tournament entry and gender, being competitive bridges 15% of the gap between choosing the least and the most prestigious tracks, and being female bridges 12.3% of that gap. To emphasize the importance of competitiveness for track choice, note that if we add all the controls from column (2) but exclude the female dummy, being competitive bridges 18% of the gap between choosing the most and the least prestigious track.[21] That is, a student's competitiveness is a

21. When running the ordered probit on tournament entry and all the basic controls (except the female dummy), $\frac{\text{Entry}}{(\text{C3}-\text{C1})}$ is 0.180 compared to $\frac{\text{Female}}{(\text{C3}-\text{C1})}$ in column (1) which is 0.154. Adding a female dummy to an OLS regression with all the controls used in column (3) increases the R^2 by .012 while adding entry raises it by .015 (adding entry on top of female raises the R^2 by a further .010 to a total of .329).

TABLE VIII
DETERMINANTS OF ACADEMIC TRACK CHOICE, INCLUDING PSYCHOLOGICAL ATTRIBUTES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Female	−0.337***	−0.272**	−0.355***	−0.299**	−0.283**	−0.230*	−0.301**	−0.256*	−0.292**	−0.252*
	(0.129)	(0.134)	(0.130)	(0.134)	(0.132)	(0.134)	(0.132)	(0.134)	(0.141)	(0.143)
Entry		0.333**		0.425***		0.336**		0.414***		0.371**
		(0.135)		(0.144)		(0.143)		(0.152)		(0.158)
Math grade	−0.094	−0.150	−0.078	−0.129	−0.113	−0.170	−0.097	−0.148	−0.133	−0.183
	(0.193)	(0.196)	(0.193)	(0.195)	(0.196)	(0.198)	(0.196)	(0.197)	(0.210)	(0.210)
GPA	0.250***	0.279***	0.248***	0.282***	0.252***	0.273***	0.251***	0.277***	0.245**	0.270***
	(0.095)	(0.095)	(0.095)	(0.095)	(0.095)	(0.095)	(0.095)	(0.095)	(0.101)	(0.101)
Math relative	−0.168	−0.187	−0.164	−0.183	−0.174	−0.189	−0.171	−0.185	−0.237	−0.249
	(0.156)	(0.157)	(0.155)	(0.156)	(0.159)	(0.160)	(0.158)	(0.158)	(0.165)	(0.165)
Math difficulty	−0.225**	−0.218**	−0.228**	−0.224**	−0.244***	−0.242***	−0.246***	−0.246***	−0.271***	−0.274***
	(0.089)	(0.089)	(0.090)	(0.090)	(0.093)	(0.092)	(0.094)	(0.094)	(0.102)	(0.101)
Math quartile	−0.329***	−0.336***	−0.334***	−0.350***	−0.338***	−0.345***	−0.343***	−0.357***	−0.338***	−0.352***
	(0.076)	(0.076)	(0.076)	(0.077)	(0.077)	(0.077)	(0.077)	(0.078)	(0.081)	(0.081)
Guessed rank			0.060	0.143*			0.061	0.131	0.031	0.090
			(0.079)	(0.083)			(0.081)	(0.086)	(0.087)	(0.091)
Risk					−0.059	−0.103	−0.049	−0.092	−0.100	−0.139*
					(0.068)	(0.069)	(0.068)	(0.069)	(0.074)	(0.075)
Lottery					0.181**	0.169**	0.181**	0.165**	0.213***	0.200***
					(0.073)	(0.074)	(0.074)	(0.074)	(0.075)	(0.075)

TABLE VIII
(CONTINUED)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Socioeconomic and age cont.									√	√
Cut 1	−0.645	−0.705	−0.375	−0.072	−0.583	−0.891	−0.253	−0.260	−4.050	−3.659
Cut 2	0.712	0.666	0.983	1.303	0.795	0.498	1.124	1.131	−2.585	−2.183
Cut 3	1.547	1.508	1.819	2.151	1.637	1.348	1.967	1.985	−1.676	−1.263
$\frac{\text{Female}}{(\text{C3}-\text{C1})}$	−0.154***	−0.123**	−0.162***	−0.134**	−0.128**	−0.103**	−0.136**	−0.114**	−0.145**	−0.126**
Diff.	20.0%		17.1%		19.4%		16.1%		12.9%	
Bootstrap p-value	0.009		0.005		0.014		0.012		0.024	
Observations	362	362	362	362	362	362	362	362	358	358

Notes. Coefficients are from ordered probit regressions, where NT > NH > ES > CS. All specifications include controls for performance in rounds 1 and 2 of the experiment, the chance of winning the round 2 tournament, school fixed effects, and test version fixed effects. The socioeconomic controls in columns (9) and (10) consist of 14 name category dummies. The age control in columns (9) and (10) is relative age measured in days. $\frac{\text{Female}}{(\text{C3}-\text{C1})}$ refers to the female coefficient divided by Cut 3 − Cut 1. Diff. refers to the percentage change of $\frac{\text{Female}}{(\text{C3}-\text{C1})}$. Robust standard errors in parentheses; p-values for $\frac{\text{Female}}{(\text{C3}-\text{C1})}$ and Diff. are bootstrapped; $^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$. The impact of confidence (comparing columns (1) and (3)) and risk attitudes (comparing columns (1) and (5)) on the gender gap ($\frac{\text{Female}}{(\text{C3}-\text{C1})}$) and the associated p-values are 5.3% (increasing) ($p = .76$) and 16.0% (decreasing) ($p = .02$), respectively.

slightly better predictor than a student's gender of study track choice.

To assess the role of competitiveness in accounting for gender differences in track choices, we compare the results of columns (1) and (2). We find a significant reduction of the effect of being female on inclusion of the tournament entry dummy. The female effect drops from 15.4% to 12.3%, a 20% reduction. This change is significant at the 1% level. This shows that the gender differences in competitiveness that have been uncovered in laboratory experiments can account for 20% of the gender gap in study track choices, even after controlling for academic performance and perceived mathematical ability.[22]

We have previously shown that tournament entry is partially explained by confidence and risk attitudes. These attributes may themselves be correlated with academic track choice. It could even be that the effect of tournament entry on track choice is driven by confidence or risk attitudes. We therefore assess the impact of competitiveness on track choice when accounting for confidence and risk attitudes. As a preliminary analysis, Figure I shows for each track the mean competitiveness of boys and girls who chose that track. In the figure, competitiveness is measured as the residual from a regression of tournament entry on the measures of performance in the experiment, the guessed rank, and the risk measures (plus school and test version fixed effects). For each gender, more competitive students select more prestigious tracks. This is a first indication that the effect of competitiveness on the study track choices of students is not due to the impact of risk attitudes and confidence alone.

In columns (3) to (8) of Table VIII, we add controls for confidence and risk attitudes to the ordered probit regressions on ranked track choice. The main result is that the coefficient of tournament entry and its effect on the gender gap in track choice remain robust and stay significant throughout. In regressions that contain both the female and the tournament entry dummy, the coefficient on entry is between 142% and 162% of the coefficient on gender. Furthermore, when having access to all the information on students as in column (7), that is objective

22. Table A.V in the Online Appendix shows that the effect of tournament entry on track choice and the drop of the female effect when the entry dummy is included are somewhat larger (20.4% and 25.9%, respectively) in a specification without objective and subjective ability variables.

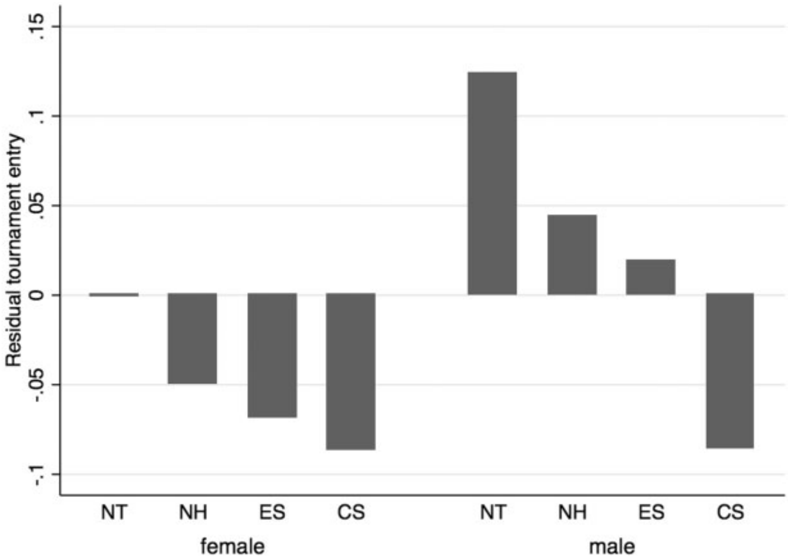

FIGURE I

Tournament Entry by Gender and Subsequent Track Choice (Conditional on Performance, Confidence, and Risk Attitudes)

The vertical axis measures the residual from a regression of tournament entry on performance on the task, confidence, and risk attitudes. The bars indicate the average values of these residuals among students that chose the different tracks by gender.

and subjective ability measures, confidence and risk measures, the tournament entry dummy is a better predictor of study track choice than is the female dummy.[23]

To confirm that competitiveness explains a substantial part of the gender gap in track choice when controlling for either confidence, risk attitudes, or both, consider pairwise comparisons between columns (3) and (4), (5) and (6), and (7) and (8). Column (2) showed that competitiveness reduces the gender gap in prestigiousness of track choice by 20% when controlling for actual and perceived academic ability. When we control for

23. When running the ordered probit on tournament entry and all the controls (except the female dummy) in column (7) of Table VIII, $\frac{\text{Entry}}{(\text{C3}-\text{C1})}$ is 0.203 compared to $\frac{\text{Female}}{(\text{C3}-\text{C1})}$ in column (7) which is –0.136. Adding a female dummy to an OLS regression with all the controls used in column (7) increases the R^2 by .009 while adding entry raises it by .016 (adding entry on top of female raises the R^2 by a further .013).

both confidence and risk-aversion measures, competitiveness still reduces the gender gap by 16% (column (7) versus column (8)). Together, competitiveness, confidence, and risk attitudes reduce the gender gap in track choice by 26% (column (1) versus (8)). Using only competitiveness, therefore, results in a reduction of the gender gap that is 78% of the size of the effect of all three psychological attributes.

In the last two columns of Table VIII we include control variables for the socioeconomic background of students based on their names (see Section II.B) and their relative age (in days). The results in columns (9) and (10) show that the impact of competitiveness remains significant.[24]

The Online Appendix shows that all of these results remain qualitatively and quantitatively similar when we use other specifications for combined profile choices, or when we use for each student their own specific ordering of prestigiousness of profiles. In the Online Appendix we also show that results remain qualitatively and quantitatively similar when instead of using ordered probit regressions on ranked track choice, we use OLS regressions. In fact, when considering how much the gender gap in track choices is reduced when controlling for the students' competitiveness, we have basically identical effects when using OLS instead of ordered probit regressions.

Instead of using confidence and risk attitudes merely as controls, we can also consider the effect of those variables on track choice and the gender gap in those choices separately. Column (3) of Table VIII shows that confidence has no significant influence on the prestige of the chosen track. Comparing columns (1) and (3) reveals that the inclusion of the confidence measure has no effect on the gender gap in choices either (which in fact increases slightly). These conclusions are unchanged when we control in addition for competitiveness (see columns (2) and (4)). Column (5) shows that risk attitudes, on the other hand, do correlate with the prestige of the chosen track. Students who opt for a more risky lottery enroll in more prestigious study tracks. Comparing columns (1) and (5) shows that adding risk attitudes reduces the gender gap by around 16% (this reduction is significant at the 5% level). The effects of competitiveness and risk attitudes on the gender gap in track choice are almost orthogonal. When we

24. For this analysis, we have to drop another four students for whom we do not have the birth date.

only control for competitiveness, we reduce the gender gap by 20% (compare columns (1) and (2)). When we only control for risk attitudes, the gender gap in choices is reduced by 16% (compare columns (1) and (5)). Adding competitiveness and risk attitudes together reduces the gender gap by 33% (compare columns (1) and (6)). That is the magnitude of the combined effect is 92% of the sum of the two separate effects. However, in contrast to competitiveness, the effect of risk attitudes on the gender gap in study track choices is somewhat variable (and not always significant) in our alternative specifications, where it ranges from 6% to 18% (see Online Appendix).

As a further analysis, we report the impact of competitiveness on selecting the most math- and science-intensive and most prestigious track, NT, compared to any other study track. Table IX reports the results from linear probability models where we regress a dummy indicating whether a student chose this profile on the same sets of controls used in Table VIII. Column (1) shows that girls are 20% less likely to choose NT than boys conditional on objective and subjective ability measures (plus school and test version fixed effects and performance in the experiment). Adding tournament entry significantly reduces this gender gap by 8.4%. This effect remains stable when we redo all the specifications from before. When we include controls for all variables—that is, when we add controls for confidence, risk attitudes, and socioeconomic background—competitiveness still reduces the gender gap in choosing NT by a significant 6.6%, see columns (9) and (10). Furthermore, depending on the specification, students who enter the tournament are between 8 and 11 percentage points more likely to choose the most prestigious profile NT. Tables A.XIV and A.XV in the Online Appendix repeat this analysis for the binary choices of a nature track versus a society track and selecting the least prestigious track, CS, compared to any other track. In Table A.XVI we consider the binary choice whether a student selected the track the student ranked highest compared to any other track.

Finally, in the Online Appendix we include tables for various robustness checks for the ordered probit results. In all of the specifications described next, the results remain qualitatively the same and stay significant. Table A.VI includes class fixed effects and Table A.IX controls for class-level characteristics (specifically the percentage of students who are female and the mean performance of students in round 2 of the experiment). Table A.VII

TABLE IX

Binary Regressions of NT versus the Rest, Including Psychological Attributes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Female	−0.195***	−0.178***	−0.202***	−0.187***	−0.189***	−0.174***	−0.196***	−0.183***	−0.178***	−0.166***
	(0.043)	(0.044)	(0.043)	(0.044)	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)	(0.046)
Entry		0.078		0.107**		0.087		0.112**		0.106*
		(0.050)		(0.054)		(0.053)		(0.056)		(0.057)
Math grade	0.007	−0.006	0.013	0.001	0.005	−0.008	0.011	−0.002	0.003	−0.009
	(0.065)	(0.065)	(0.065)	(0.065)	(0.065)	(0.066)	(0.066)	(0.066)	(0.067)	(0.068)
GPA	0.024	0.031	0.023	0.031	0.024	0.029	0.023	0.030	0.021	0.028
	(0.031)	(0.032)	(0.031)	(0.032)	(0.031)	(0.032)	(0.032)	(0.032)	(0.033)	(0.034)
Math relative	−0.057	−0.060	−0.055	−0.059	−0.057	−0.059	−0.056	−0.058	−0.061	−0.063
	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.055)	(0.055)
Math difficulty	−0.068**	−0.067**	−0.069**	−0.068**	−0.071**	−0.070**	−0.071**	−0.071**	−0.074**	−0.075**
	(0.032)	(0.032)	(0.032)	(0.032)	(0.033)	(0.032)	(0.033)	(0.033)	(0.034)	(0.034)
Math quartile	−0.086***	−0.087***	−0.088***	−0.091***	−0.087***	−0.088***	−0.089***	−0.091***	−0.087***	−0.089***
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.027)	(0.027)
Guessed rank			0.023	0.044			0.023	0.042	0.021	0.039
			(0.026)	(0.028)			(0.027)	(0.029)	(0.028)	(0.029)
Risk					−0.009	−0.021	−0.005	−0.017	−0.012	−0.023
					(0.024)	(0.024)	(0.025)	(0.025)	(0.025)	(0.025)
Lottery					0.020	0.017	0.020	0.015	0.024	0.021
					(0.024)	(0.024)	(0.024)	(0.024)	(0.023)	(0.023)
Socioeconomic and age cont.									√	√
Diff.	8.4%		7.6%		7.7%		6.7%		6.6%	
Bootstrap p-value	0.060		0.025		0.052		0.027		0.034	
Observations	362	362	362	362	362	362	362	362	358	358

Notes. Coefficients are from OLS regressions, where the outcome variable is a dummy indicating a choice of NT. All specifications include controls for performance in rounds 1 and 2 of the experiment, the chance of winning the round 2 tournament, school fixed effects and test version fixed effects. The socioeconomic controls in columns (9) and (10) consist of 14 name category dummies. The age control in columns (9) and (10) is relative age measured in days. Diff. refers to the percentage change of female. Robust standard errors in parentheses; p-values for Diff. are bootstrapped. $^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$. The impact of confidence (comparing columns (1) and (3)) and risk attitudes (comparing columns (1) and (5)) on the gender gap (Female) and the associated p-values are 3.5% (increasing) ($p = .240$) and 3.1% (decreasing) ($p = .023$), respectively.

reports the full table with name category dummies and the age control. In Table A.VIII, we run weighted ordered probit regressions to test whether it matters for our results that the track choices of students in our sample differ slightly from the national average. As already discussed, the Online Appendix also contains a number of specifications where we treat combi-track students in different ways. In Table A.X, we treat ES/CS students as ES and NH/NT students as NT. In Table A.XI, we treat these tracks as separate (using NT > NT/NH > NH > ES > ES/CS > CS as our ranking). In Table A.XII we repeat the analysis where instead of ranking tracks by prestigiousness and math and science intensity, we use for each student the rank the student herself gave to the track she chose. Finally in Table A.XVII we present results from OLS regressions instead of ordered probit models, where we replicate the analyses reported in Table VIII.

VII. Discussion

We showed the importance of competitiveness as measured by an experiment where students can choose to enter a tournament for explaining study track choices and the gender difference in those choices. The results were robust even after controlling for confidence, risk attitudes, and socioeconomic status. In this section, we address whether we can attribute the effect of our laboratory measure of tournament entry to competitiveness. In what follows, we discuss three alternative interpretations of our results. Tournament entry conditional on performance may be an (additional) measure of the students' perceived mathematical ability, their actual mathematical ability, or their preference for math.

Concerning tournament entry being an additional measure for perceived math ability, note that when we add tournament entry to ordered probit regressions on study track choice, the effects of perceived math ability are not substantially altered (compare columns (1) and (2) in Table VIII). Furthermore, the guessed rank in the experimental task, which might be an even better additional predictor of a student's perceived math ability, is not correlated with the study track choice, despite being significantly correlated with the decision to enter the tournament.

To assess whether the decision to enter the tournament is mostly an additional measure of actual math ability, note that all regressions control for performance in the experiment. In addition, when we add entry to ordered probit regressions on study track choice, the effects of actual math ability are not substantially altered (again compare columns (1) and (2) in Table VIII). Furthermore, conditional on subjective math ability the absolute and relative math grades do not significantly predict study track choice (column (1) in Table VIII: absolute and relative math grades are not jointly significant [$p = .39$]). The coefficient on tournament entry, however, is both significant and much larger than the coefficient on the standardized math grade. Finally, if anything, the coefficient on GPA slightly increases rather than decreases when we add the entry decision in Table VIII.[25]

Finally, the decision to enter the tournament could be a measure of the students' preference or tolerance for math. One channel could be that a preference for math translates to more optimistic beliefs on relative performance in math tasks, which we already discussed. A second possible mechanism is that a preference for math translates into a preference or tolerance for feedback on relative performance or a relative payment scheme. Although we have no direct evidence addressing this concern in the present article, the results of Niederle and Vesterlund (2007) suggest that this is not the case. Specifically, they constructed a second choice environment where subjects could receive information about their relative performance and be paid depending on their relative performance without having to compete. Their paper found no gender differences in choices of such a relative payment scheme over a piece rate payment.[26] Overall, we are

25. To be precise, in column (3) of Table A.V in the Online Appendix we find that a one standard deviation increase in GPA bridges 12.3% of the gap between choosing the least and the most prestigious track ($\frac{\text{GPA}}{(\text{Cut3}-\text{Cut1})}$). In column (4), when we add entry, the effect of the GPA is 13.6%. Similar results hold for comparing the impact of an increase in the GPA between columns (7) and (8), columns (9) and (10), columns (11) and (12), and columns (13) and (14).

26. Specifically, after performing in a piece rate, then a tournament, and then a treatment where subjects chose between those two payment schemes (as in our own experiment), subjects encountered the following choice in the fourth and last round. In case this round was chosen for payment, participants were paid according to their round 1 piece rate performance but had to decide how to be paid: either via a piece rate or via a tournament, where the person with the highest round 1 piece rate performance wins. The choice of tournament in this "submit the piece rate" round (which is identical to the round 3 tournament entry choice except for the need to perform and compete) showed no gender difference once the round 1 piece rate performance and beliefs on the relative piece rate performance were controlled for.

confident that the choice to enter a tournament is a measure of competitiveness rather than of actual or perceived math performance or a preference for math.

VIII. Conclusion

In this article we presented evidence that an incentivized measure of competitiveness is a relevant predictor of important education choices of young people in the Netherlands. More important, we showed that a substantial share of gender differences in these education choices can be attributed to gender differences in competitiveness as measured by our experiment.

This article is part of a small but growing literature that aims to predict economic outcomes outside of the laboratory with laboratory measures, see, for example, Karlan (2005), Ashraf, Karlan, and Yin (2006), Fehr and Goette (2007), Meier and Sprenger (2010), Dohmen et al. (2011), Dohmen and Falk (2011), and Zhang (2012a). This is a promising and important approach to show the external validity of traits measured in the lab, but more important to show their external relevance. One main challenge in this line of research is to beware of reverse causality. This would, for example, have been a significant concern had we measured competitiveness after students made their choices and when they all have different classroom experiences. This is why we administered the experiment while students still shared the same experiences, several months before they made their education choice.

The paper most closely related to ours that studies the external relevance of competitiveness is Zhang (2012a). She conducts a standard Niederle and Vesterlund (2007) competitiveness experiment with middle schoolers from Ninglang County in China and observes their decision to take a very competitive entry exam for high school. She finds that students more inclined to compete are more likely to take the entry exam, controlling for the test score on a previous exam. The results indicate no large gender difference in either take-up rates of the entry exam or, perhaps more surprising, in tournament entry. The latter is in contrast to other studies that found gender differences in competitiveness among children (Sutter and Rützler, 2010), or Zhang (2012b) who finds gender differences for ethnic minorities among high school children from the same area.

A second approach for testing the external validity of laboratory results is to mimic laboratory experiments in a richer and ideally more naturally occurring field setting. Although gender differences in competitiveness have been repeatedly documented in the field, there has only been limited evidence on gender differences in tournament entry (see Niederle and Vesterlund 2011 for a survey). In a recent study, Flory, Leibbrandt, and List (2010) conducted a field experiment in which job-seekers were randomly offered compensation schemes that varied in the degree of competition. In accordance with the findings from the laboratory, they find that women are relatively less likely to apply for a job with a competitive payment scheme than are men. While it is reassuring that both the intensive and extensive margin of gender differences in competitiveness can be found in additional specific groups beyond school and college students, such evidence does not directly inform us whether gender differences in competitiveness can account for an economically significant portion of observed gender differences in educational choices and labor market outcomes. That is, although the external validity of laboratory results to other subject pools are confirmed, such studies do not necessarily address the external relevance of the concept to standard economic questions. Such standard field experiments are therefore probably better thought of as complements to laboratory experiments rather than substitutes to research addressing the external relevance of experimental variables.

By validating the importance of competitiveness, our article opens up new research questions. For example, how does competitiveness predict the performance of students in various study tracks? One could imagine that competitive students fare better in terms of grades than do their less competitive peers. On the other hand, competitiveness may lead students to "overreach" and enter study tracks that are too difficult for them. We saw that especially some boys aim for the most mathematically heavy NT track while scoring high on competitiveness but not so much on the math grade. It will be important to understand the extent to which competitiveness affects the study track choices of students, and the extent to which it affects the performance of students once they choose certain tracks. If competitiveness mostly affects the choices, this suggests the policy-relevant question of whether different choice environments can affect study choices because different choice environments reward different psychological attributes such as competitiveness (see, e.g., Niederle and Yestrumskas 2008).

Future research will determine whether our results can be replicated in other environments and with different or larger subject pools. In our environment, prestige and math intensity are very correlated, and it remains to be determined whether our results hold when this is not the case. Likewise, it remains to be determined to what extent a competitiveness measure on other, less math-oriented tasks, correlates with the present measure and with study track choices.

Another question concerns whether competitiveness influences other economic decisions or only the choice of study tracks in secondary school. Using a representative sample of higher educated men and women in the Netherlands, Kalmijn and Van der Lippe (1997) estimate that almost 10% of the gender wage gap of 0.33 can be attributed to gender differences in fields of study. If gender differences in competitiveness would only operate through the academic tracks studied here, this suggests that they explain 2% of the gender wage gap. Gender differences in competitiveness may, however, also influence the wage gap through other channels and may potentially explain a share of the part that is unexplained by observable characteristics. By following the students of our sample into the labor market, we plan to report on this in future work.

A final set of open questions concerns the trait of competitiveness. What does the choice of entering a tournament exactly measure, and how is this measure correlated with other traits that may potentially be more familiar but could be hard to capture? For example, how does competitiveness differ from traits like ambition or challenge seeking? Which psychological traits correlate with competitiveness? Finally, an important open question is whether we can manipulate the competitiveness of students and whether this would affect their educational choices.

University of Amsterdam
Stanford University and National Bureau of Economic Research
University of Amsterdam

Supplementary Material

An Online Appendix for this article can be found at QJE online (qje.oxfordjournals.org).

References

Altonji, Joseph G., and Rebecca Blank, "Race and Gender in the Labor Market," in *Handbook of Labor Economics*, Orley C. Ashenfelter, and David Card, eds. vol. 3C (Amsterdam: Elsevier Science Publishers, 1999), 3144–3259.

Ashraf, Nava, Dean Karlan, and Wesley Yin, "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines," *Quarterly Journal of Economics*, 121 (2006), 635–672.

Balafoutas, Loukas, and Matthias Sutter, "Affirmative Action Policies Promote Women and Do Not Harm Efficiency in the Laboratory," *Science*, 335, no. 6068 (2012), 579–582.

Bertrand, Marianne, "New Perspectives on Gender," in *Handbook of Labor Economics*, Orley Ashenfelter, and David Card, eds. vol. 4B (Amsterdam: Elsevier, 2011), 1543–1590.

Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz, "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors," *American Economic Journal: Applied Economics*, 2 (2010), 228–255.

Bloothooft, Gerrit, and David Onland, "Socioeconomic Determinants of First Names," *Names*, 59 (2011), 25–41.

Brown, Charles, and Mary Corcoran, "Sex-Based Differences in School Content and the Male-Female Wage Gap," *Journal of Labor Economics*, 15 (1997), 431–465.

CBS, *Jaarboek onderwijs in cijfers 2010* (Centraal Bureau voor de Statistiek, 2010), http://www.cbs.nl/nl-NL/menu/themas/onderwijs/publicaties/publicaties/archief/2010/2010-f162-pub.htm.

———, *Jaarboek onderwijs in cijfers 2012* (Centraal Bureau voor de Statistiek, 2012), http://www.cbs.nl/nl-NL/menu/themas/onderwijs/publicaties/publicaties/archief/2012/2012-f162-pub.htm.

Croson, Rachel, and Uri Gneezy, "Gender Differences in Preferences," *Journal of Economic Literature*, 47 (2009), 448–474.

Dohmen, Thomas, and Armin Falk, "Performance Pay and Multidimensional Sorting: Productivity, Preferences, and Gender," *American Economic Review*, 101 (2011), 556–590.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jurgen Schupp, and Gert G. Wagner, "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences," *Journal of the European Economic Association*, 9 (2011), 522–550.

Dreber, Anna, Emma von Essen, and Eva Ranehill, "Gender and competition in adolescence: task matters," *Experimental Economics*, 17 (2014), 154–172.

Eccles, Jacquelynne, "Perceived Control and the Development of Academic Motivation: Commentary," *Monographs of the Society for Research in Child Development*, 63 (1998), 221–231.

Eckel, Catherine C., and Philip J. Grossman, "Sex Differences and Statistical Stereotyping in Attitudes toward Financial Risk," *Evolution and Human Behavior*, 23, no. 4 (2002), 281–295.

Ellison, Glenn, and Ashley Swanson, "The Gender Gap in Secondary School Mathematics at High Achievement Levels: Evidence from the American Mathematics Competitions," *Journal of Economic Perspectives*, 24, no. 2 (2010), 109–128.

Fehr, Ernst, and Lorenz Goette, "Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment," *American Economic Review*, 97 (2007), 298–317.

Felder, Richard, Gary Felder, Merdith Mauney, Charles Hamrin, and Jacquelin Dietz, "A Longitudinal Study of Engineering Student Performance and Retention. III. Gender Differences in Student Performance and Attitudes," *Journal of Engineering Education*, 84, no. 2 (1995), 151–163.

Flory, Jeffrey A., Andreas Leibbrandt, and John A. List, "Do Competitive Work Places Deter Female Workers? A Large-Scale Natural Field Experiment on Gender Differences in Job-Entry Decisions," NBER Working Paper No. w16546 (2010).

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini, “Performance In Competitive Environments: Gender Differences,” *Quarterly Journal of Economics*, 118 (2003), 1049–1074.
Goldin, Claudia, Lawrence F. Katz, and Ilyana Kuziemko, “The Homecoming of American College Women: The Reversal of the College Gender Gap,” *Journal of Economic Perspectives*, 20, no. 4 (2006), 133–156.
Goodman Research Group, “Final Report of the Women’s Experience in College Engineering Project,” report (2002), http://www.grginc.com/WECE_FINAL_REPORT.pdf.
Grogger, Jeff, and Eric Eide, “Changes in College Skills and the Rise in the College Wage Premium,” *Journal of Human Resources*, 30 (1995), 280–310.
Huguet, Pascal, and Isabelle Regner, “Stereotype Threat among School Girls in Quasi-Ordinary Classroom Circumstances,” *Journal of Educational Psychology*, 99 (2007), 545–560.
Kalmijn, Matthijs, and Tanja Van der Lippe, “Type of Schooling and Sex Differences in Earnings in the Netherlands,” *European Sociological Review*, 13 (1997), 1–15.
Kamas, Linda, and Anne Preston, “Are Women Really Less Competitive Than Men?,” unpublished manuscript, Santa Clara University (2010).
Karlan, Dean, “Using Experimental Economics to Measure Social Capital and Predict Financial Decisions,” *American Economic Review*, 95 (2005), 1688–1699.
LeFevre, Jo-Anne, Alison G. Kulak, and Stephanie L. Heymans, “Factors Influencing the Selection of University Majors Varying in Mathematical Content,” *Canadian Journal of Behavioural Science*, 24, no. 3 (1992), 276–289.
Lonnqvist, Jan-Erik, Markku Verkasalo, Gari Walkowitz, and Philipp C. Wichardt, “Measuring Individual Risk Attitudes in the Lab: Task or Ask? An Empirical Comparison,” technical report, DIW Berlin (2010).
Meier, Stephan, and Charles Sprenger, “Present-Biased Preferences and Credit Card Borrowing,” *American Economic Journal: Applied Economics*, 2 (2010), 193–210.
Murnane, Richard J., John B. Willett, Yves Duhaldeborde, and John H. Tyler, “How Important Are the Cognitive Skills of Teenagers in Predicting Subsequent Earnings?,” *Journal of Policy Analysis and Management*, 19 (2000), 547–568.
Niederle, Muriel, Carmit Segal, and Lise Vesterlund, “How Costly Is Diversity? Affirmative Action in Light of Gender Differences in Competitiveness,” *Management Science*, 59 (2013), 1–16.
Niederle, Muriel, and Lise Vesterlund, “Do Women Shy Away from Competition? Do Men Compete Too Much?,” *Quarterly Journal of Economics*, 122 (2007), 1067–1101.
———, “Gender and Competition,” *Annual Review of Economics*, 3 (2011), 601–630.
Niederle, Muriel, and Alexandra H. Yestrumskas, “Gender Differences in Seeking Challenges: The Role of Institutions,” Unpublished manuscript (2008).
Paglin, Morton, and Anthony M. Rufolo, “Heterogeneous Human Capital, Occupational Choice, and Male-Female Earnings Differences,” *Journal of Labor Economics*, 8 (1990), 123–144.
Pautler, Emmanuelle, “The Links between Secondary and Higher Education in France,” *European Journal of Education*, 16, no. 2 (1981), 185–195.
Preckel, Franzis, Thomas Goetz, Reinhard Pekrun, and Michael Kleine, “Gender Differences in Gifted and Average-Ability Students,” *Gifted Child Quarterly*, 52, no. 2 (2008), 146–159.
Roeder, Peter M., and Sabine Gruehn, “Geschlecht und Kurswahlverhalten,” *Zeitschrift für Pädagogik*, 43 (1997), 877–894.
Schroter Joensen, Juanna, and Helena Skyt Nielsen, “Is There a Causal Effect of High School Math on Labor Market Outcomes?,” *Journal of Human Resources*, 44 (2009), 171–198.
———, “Math and Gender: What if Girls Do Math?,” Unpublished manuscript, Stockholm School of Economics (2011).

Sutter, Matthias, and Daniela Rützler, "Gender Differences in Competition Emerge Early in Life," IZA Discussion Paper No. 5015 (2010).
Weinberger, Catherine J., "Mathematical College Majors and the Gender Gap in Wages," *Industrial Relations: A Journal of Economy and Society*, 38 (1999), 407–413.
———, "Is Teaching More Girls More Math the Key to Higher Wages?," in *Squaring Up: Policy Strategies to Raise Women's Incomes in the United States*, Mary C. King, ed. (Ann Arbor: University of Michigan Press, 2001), 226–258.
———, "Is the Science and Engineering Workforce Drawn from the Far Upper Tail of the Math Ability Distribution?," Unpublished manuscript (2005).
Wozniak, David, William T. Harbaugh, and Ulrich Mayr, "The Menstrual Cycle and Performance Feedback Alter Gender Differences in Competitive Choices," *Journal of Labor Economics*, 32 (2014), 161–198.
Zhang, Y. Jane, "Can Experimental Economics Explain Competitive Behavior Outside the Lab?," Unpublished manuscript (2012a).
———, "The Communist Experiment in China: Narrowing the Gender Gap in Competitive Inclination," Unpublished manuscript (2012b).

This page intentionally left blank