Gender, Competitiveness and Career Choices*

Thomas Buser
University of Amsterdam and *TIER*

Muriel Niederle
Stanford University and *NBER*

Hessel Oosterbeek
University of Amsterdam and *TIER*

October 15, 2012

Abstract

Gender differences in competitiveness are often discussed as a potential explanation for gender differences in labor market outcomes. We correlate an incentivized measure of competitiveness with the first important career choice of secondary school students in the Netherlands. At the age of 15, these students have to pick one out of four study profiles, which vary in how prestigious they are. While boys and girls have very similar levels of academic ability, boys are substantially more likely than girls to choose more prestigious profiles. We find that up to 23 percent of this gender difference can be attributed to gender differences in competitiveness. This lends support to the extrapolation of laboratory findings on competitiveness to labor market settings.

JEL-codes: C9, I20, J24, J16
Keywords: field and laboratory experiments, economics of education, competitiveness

1 Introduction

Gender differences in education and labor market outcomes, though greatly reduced, have remained ubiquitous. To understand gender differences in these outcomes, psychological attributes are commonly discussed as potential explanations. While the last decade saw a flurry of laboratory research documenting gender differences in psychological attributes such as competitiveness, there has been no satisfactory direct evidence linking them to education and labor market outcomes.[1] This is exactly what this paper does. We focus on competitiveness, which is an attribute for which large

*We thank the staff of the four schools that allowed us to collect the data from their pupils that we use in this paper. 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.

[1] Bertrand (2011) summarizes this literature and concludes: "While the laboratory evidence shows in many cases large gender differences (say, in attitudes towards risk, or attitudes toward competition), most of the existing attempts to measure the impact of these factors on actual outcomes fail to find large effects. This is undoubtedly a reflection of a rather new research agenda, as well as of the difficulty in finding databases that combine good measures of psychological attributes with real outcomes. More direct demonstrations of field relevance will be crucial for these new perspectives to have a lasting impact on how labor economists approach their study of gender gaps" (p.1583).

gender differences have been documented (see Croson and Gneezy, 2009 and Niederle and Vesterlund, 2011). Through in-class experiments we collected data on the competitiveness of high school students which we merged with information about their subsequent education choices. With this data we demonstrate that competitiveness significantly correlates with educational choices and explains an economically and statistically significant part of the gender gap in those choices.

Gender differences in educational choices, particularly regarding intensity in math and science subjects, remain significant. While in the U.S. girls take on average as many advanced math and science classes as boys and perform on average at similar levels (Goldin et al., 2006), this is not the case in many OECD countries.[2] Even in the U.S., girls are underrepresented among extremely high achieving math students (Ellison and Swanson, 2010). At the college level, in the U.S., women are significantly less likely than men to graduate from a major in science, technology, engineering or mathematics.[3] The reason to be concerned about gender differences in math and sciences compared to, say, literature, is that the choices of math and science classes are most predictive of college attendance and completion (Goldin et al., 2006). Furthermore, performance in mathematics has consistently been found to serve as a predictor for future earnings. For example, Paglin and Rufolo (1990) report that a large fraction of the gender gap in average starting salaries for college graduates is between rather than within detailed college major (for additional evidence and discussion see Grogger and Eide, 1995; Brown and Corcoran, 1997; Weinberger, 1999; Weinberger, 2001; Murnane et al., 2000; Altonji and Blank, 1999).[4]

Next to discrimination (Goldin and Rouse, 2000) and differences in preferences (which could be driven by stereotypes), a standard explanation for gender differences in education choices is differences in ability.[5] However, Ellison and Swanson (2010) provide compelling evidence that the gender imbalance in the U.S. among high achieving math students is not driven solely by differences in mathematical ability. They show that in mathematics, high-achieving boys come from a variety of backgrounds as would be expected from an allocation of talent that shows some randomness, while high-achieving girls are almost exclusively drawn from a small set of super-elite schools. Furthermore, research investigating career choices of women and men suggests that among equally gifted students, males are much more likely to choose a math heavy college major (see LeFevre et al., 1992; Weinberger, 2005).

In line with the recognition of the importance of non-cognitive skills for educational and labor

[2] We will show that in the Netherlands boys are significantly more likely to take math classes in high school than girls. In France, where like in the Netherlands high school children decide on which sets of classes to enroll in, girls are less likely to choose the math and science heavy options (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).

[3] http://nces.ed.gov/pubs2009/2009161.pdf.

[4] In a study on the gender gap in earnings among MBA's from Chicago Booth, Bertrand et al. (2011) conclude that one of three factors that account for the large gender gap in earnings a decade after MBA completion is differences in training prior to MBA graduation, with, most notably, women taking many fewer finance courses than men.

[5] For the potential importance of stereotypes on preferences of females over mathematics see Nosek et al. (2002), Kiefer and Sekaquaptewa (2007) and Ceci et al. (2009). For evidence of the presence of such stereotypes already in elementary school see Cvencek et al. (2011).

market outcomes (Cunha and Heckman, 2007; Segal, 2012; Borghans et al., 2008) a recent literature in behavioral economics has focused on gender differences in psychological attributes as another potential source for the observed gender differences. One large and robust gender difference is that women are found to be less competitive than men (see Gneezy et al., 2003, and, for an overview on gender differences Croson and Gneezy, 2009). In the first paper on gender differences in the willingness to compete, Niederle and Vesterlund (2007) assess choices of college students who perform equally well in a simple arithmetic task. They find that while 73 percent of men choose a competitive tournament payment scheme instead of a non-competitive piece-rate compensation, only 35 percent of women do so. Subsequent research has confirmed this gender difference in the willingness to compete (see Niederle and Vesterlund, 2011).

It seems plausible that competitiveness is important for educational choices and labor market outcomes. People who shy away from competitive environments may self-select into different, potentially lower paid, careers (Kleinjans, 2009). The fact that women shy away from competition more than men may also account for the fact that few qualified women reach the top (see Bertrand and Hallock, 2001) and for the acceleration of the gender wage gap in the upper tail (see Albrecht et al., 2003; Arulampalam et al., 2007). Furthermore, competitiveness could be an especially important trait for certain fields such as sciences and mathematics which are viewed as competitive. One reason may be that it is easier to rank answers in math tests than in say verbal tests.[6] Furthermore, if more boys select math heavy courses and majors, this increases the number of potential male competitors. Experiments have shown that for women both the performance in (Gneezy et al., 2003) as well as the selection into competitive environments (Niederle et al., 2012; Balafoutas and Sutter, 2012) is sensitive to the gender composition of the group.[7] There is also evidence of low tolerance for competition among women who drop out of math intensive college majors and engineering.[8] However, most of this evidence is casual and may suffer from the problem of reverse causality. Women who hold low-profile jobs may not encounter many competitive situations and potentially become less inclined towards competitions, or may simply start to describe themselves as less competitive given the job they hold. Similarly, women who drop out of science and engineering may search for explanations such as the negative aspect of the competitive environment.

To assess the effect of competitiveness on educational choices, we therefore aim to measure competitiveness before students have different experiences resulting from their choices, as these different experiences could affect their measure of competitiveness. We run our study in the Netherlands where, at the end of the third year of secondary school, students in the pre-university track choose between four study profiles which strongly correlate with the choice of major in tertiary education

[6] Indeed, laboratory research has shown that gender differences in competitiveness are sometimes (e.g. Kamas and Preston, 2010) but not always (e,g, Wozniak et al., 2010; see Niederle and Vesterlund, 2011 for an overview) attenuated when assessed in verbal tasks.

[7] 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.

[8] The report "Women's Experiences in College Engineering" (Goodman Research Group, 2002) reports that women do not drop their math intensive studies because of ability, but rather low self-confidence. These women also mention negative aspects of their climate such as competition and discouraging faculty and peers (see also Felder et al., 1995, for a study on engineering).

and labor market outcomes. The four profiles, science, health, social sciences and humanities, are clearly ranked not only in terms of their math intensity, but also in terms of academic prestige, in exactly that order. Girls, despite being slightly better academically than boys, are less likely to enroll into the most prestigious and most math heavy science profile and more likely to enroll in the least prestigious humanities profile.

We administered an experiment in four schools in and around Amsterdam just before students chose their study profiles. Since we are concerned with the choice of prestigious profiles typically favored by males we measure the competitiveness of students in a stereotypical male task. Specifically, we use the most common measure of competitiveness by Niederle and Vesterlund (2007), which has proven to be robust across many settings and subject pools (Niederle and Vesterlund, 2011). Niederle and Vesterlund (2007) show that gender differences in competitiveness can be partially attributed to gender differences in confidence, while gender differences in risk attitudes play only a minor role. Both confidence and risk attitudes could also play an important role when students decide whether to choose a more prestigious study profile. We therefore also administer incentivized measures of students' confidence and their risk attitudes. The schools provided us with the subsequent profile choices of students as well as with their grades. Finally, we assess the students' perceptions of their mathematical ability, as grades may potentially not be the most accurate predictor of ability.

Students in our sample exhibit the expected significant gender gap in prestigiousness (or math intensity) of chosen study profile, controlling for both objective and subjective academic performance. Confirming the results from the literature, boys in our sample are also more than twice as likely than girls to enter the tournament. The first result is that competitiveness correlates positively with the prestige of chosen study profiles. Being competitive bridges around 20 percent of the distance between choosing the lowest and the highest ranked profile, which is comparable to the effect of being female. Our main finding is that competitiveness accounts for 23 percent of the gender gap in the prestige of chosen study profiles. When we control for grades and perceived mathematical ability, this percentage equals 18. When we subsequently also control for confidence and risk attitudes, inclusion of competitiveness closes the gender gap in the prestige of chosen profiles by a significant 15 percent.

This paper is related to a growing literature on the external validity of lab results on gender differences, most notably competitiveness. Typically, such field evidence aims to either find or create environments that resemble the experimental design, albeit with other subject pools. While the field evidence has focused on showing that the gender gap in performance increases with rising competitive pressure, the scant early studies on choices of competitive environments find evidence consistent with women shying away from competition (see e.g. Flory et al., 2010, and for an overview Niederle and Vesterlund, 2011). While it is reassuring that gender differences in competitiveness can be found in additional specific groups beyond school and college students, such evidence does not 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, the external *relevance* of the concept of competitiveness has not yet been addressed in a satisfactory manner. Our paper fills that gap. Our results show that the competitiveness measure commonly used in laboratory experiments helps uncover a trait which accounts for a statistically and economically significant portion of the gender difference in educational choices. As such our paper not only shows the external relevance of competitiveness but also validates the specific measure of competitiveness provided by Niederle and Vesterlund (2007).

The remainder of this paper is organized as follows. Section 2 describes the data collection. Section 3 provides details and results concerning the education choice we analyze in this paper, and its context in the Dutch education system. Section 4 describes and analyzes the experimental data. Section 5 presents and discusses the main results. Section 6 summarizes and concludes.

2 Study design

2.1 Data collection

The students participating in this study are drawn from the population of Dutch secondary school students who are enrolled in the pre-university track. Halfway through the six years of secondary school, at the end of grade 9, students in the pre-university track have to choose one out of four study profiles: the science-oriented profile Nature & Technology (NT), the health-oriented profile Nature & Health (NH), the social science-oriented profile Economics & Society (ES) and the humanities-oriented profile Culture & Society (CS).

We invited secondary schools in and around Amsterdam to participate in a research project investigating the determinants of study profile choices. We demanded one class hour (45 or 50 minutes) of all grade 9 classes in the pre-university track. The invitation letter stated that students would participate in an in-class experiment and be paid depending on their choices. It also mentioned that after the experiment a short questionnaire would be administered. For detailed instructions see the online appendix.

Four schools cooperated, one in the city of Amsterdam and three in cities close to Amsterdam. In each school, we captured all students in the 3rd grade of the pre-university track. A total of 397 students in 16 classes participated. Because the schools are geographically dispersed, we do not worry that students receive 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 of 2011.

After the end of the school year, the schools provided us with the final grades and the definite profile choices of the students. For 35 students we do not have such a definite profile choice.[9] For 20 of these students, we can use information about their expected profile choice obtained through the short questionnaire.[10] We drop the remaining 15 students for whom we have neither a definite choice nor a clear choice from the questionnaire. We have to drop an additional 4 students from the

[9]Some students may have to retake the year, and in some schools those are included in the final profile choice, in others not.

[10]For the students for whom we have both the definite profile choice and the intention stated in the questionnaire, the questionnaire answer accurately predicts the final choice in 93 percent of the cases.

analysis because they showed up late to class and missed part of the experiment, 2 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.

2.2 Variables

Competitiveness. We use a classroom experiment to obtain an individual measure of competitiveness. The design closely follows Niederle and Vesterlund (2007). Participants perform a real task, first under a noncompetitive piece rate scheme and then under a competitive tournament scheme. Participants choose which of the two payment schemes to apply to their third and final performance. This allows us to determine the extent to which the choice of compensation scheme depends on performance.

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. 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. At the end of three minutes subjects had to drop the pen and stand up. In each round there were three versions of the 26 addition problems to prevent copying from neighbors.[11]

Participants were informed that they would perform in three rounds, one of which was randomly chosen for payment at the end of the experiment through the roll of a die in front of the class. Participants received details on each round only immediately before performing in the task. Participants did not receive any information about their own performance or the performance of others. They were paid a week later through sealed envelopes, at which time they could make inferences about their relative performance. Participants earned on average €5.55, with a minimum of zero and a maximum of €25 (this includes the payment from incentivized questions to elicit confidence and risk attitudes; see below).

Participants first performed the task under a noncompetitive piece rate of 25 Euro-cents per correctly solved problem. In round two they performed in tournaments of four, where the three 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 would be paid €1 per correct problem and the others received no payment. In case of a tie, the winner was randomly determined.

In the third round, participants chose which of the two payment schemes they would prefer. Students were informed that in case round three was selected for payment, the earnings were computed as follows. A participant who chose the piece rate received 25 cents per correct problem. A participant who selected the tournament would win if his or her new round 3 performance exceeded the performance of the other three group members in the previous round 2 tournament. Therefore,

[11] Since including controls for the specific test does not change any result, we omit them for brevity.

just like in Niederle and Vesterlund (2007), the choice was an individual decision as a subject could not affect the payoffs of any other participant.[12]

Confidence. To measure confidence we ask participants for their beliefs about their relative performance after completing Round 3. Specifically, we asked students about their relative performance in the Round 2 tournament compared to the other three group members, from 1 (best) to 4 (worst) of their group of four. If their guess was correct, they received €1.[13]

Risk attitudes. We elicit risk attitudes by using two measures. First, following Eckel and Grossman (2002), subjects picked one option among a sure payoff of €2 and four 50/50 lotteries in Euros 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 is determined by a dice roll 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 try to avoid taking risks?" The answer is on a scale from 0 ("unwilling to take risks") to 10 ("fully prepared to take risk"). This second risk measure, being a survey question, is cheap, but potentially less reliable. However, Dohmen et al. (2011), using representative survey data from Germany, find that this simple question predicts both choices in a lottery task and risky behavior 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 for risk attitudes.

Ability. We use information about students' grades at the end of 9^{th} grade provided by the schools to construct three objective ability measures. The first is GPA, the second the grade for mathematics. In the Dutch school system, grades are expressed on a scale from 1(worst) to 10(best), where 6 is the first passing grade. The third measure is the relative math grade in the student's class. We gave the best students in class a rank of 1. The rank of each student is equal to 1 plus the number of students with a strictly better grade. We then normalize the measure by dividing by the number of students in the class.[14]

Since grades need not be a good predictor of mathematical ability, we collected subjective ability measures in the questionnaire that was administered after the in-class experiment. We asked 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%).[15] We also asked students how difficult

[12] 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 provides no externality to another subject, hence motives such as altruism, or fear of interfering with someone else's payoff play no role.

[13] 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.

[14] To compute the relative mathematics rank in class, we include all 397 students in our sample, including the 35 students we had to drop for the final results.

[15] This was phrased as three yes/no questions: "Do you think your mathematics ability is in the top 25% of your

they find it to pass their math class on a scale from 0 (very easy) to 10 (very hard). While these questions may yield a better assessment of mathematical ability, they could in addition be a measure of confidence, which in turn could influence study profile 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). This may in part be driven by gender stereotypes that boys are better at math than girls and by the fact that the fraction of male teachers is larger in courses that are math intensive.[16]

We present the results from the study in three stages. First, we describe the environment and study profile decisions of students. We document significant gender differences in the prestige (and math intensity) of the profile choices which are not explained by differences in ability. Second, we present the data on the experimentally measured attributes: competitiveness, confidence and risk aversion. We show significant gender differences in competitiveness, that is selection into the tournament conditional on performance. We assess the extent to which these differences can be attributed to gender differences in confidence and risk attitudes. Finally, in the main result section we examine whether competitiveness correlates with profile choice. We then assess to what extent gender differences in competitiveness can account for gender differences in the prestige of chosen profiles.

3 The Education Choice

3.1 Study Profiles

The students participating in this study are drawn from the population of Dutch secondary school students who are enrolled in the pre-university track. 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 tracks: around 20 percent of students graduate from the six-year pre-university track, 25 percent from the five-year general track and 55 percent from the four-year vocational track. Who enrolls in which track is to a large extent determined by the score on a nation-wide achievement test administered at the end of primary school. Girls are somewhat more likely to go to the pre-university track, making up 54 percent of the students (Statistics Netherlands).

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

- the science-oriented profile Nature & Technology (NT)

year?", "...top 50% of your year?", "...top 75% of your year?". A student who answered all 3 questions with a 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%). 67 students, however, switched back to no. For these students, we count the first yes as their true answer.

[16] Dee (2007) and Carrell et al. (2010) show that having a female teacher increases the math and science performance of female students, where the effect is particularly large at the upper tail. However, teachers in math are less likely to be female than those in other subjects. Using the 1999-2000 Schools and Staffing Survey (SASS), Dee (2007) estimates that in 12th grade 44 and 52 percent of science and math teachers are female, compared to 71 percent in reading. Bettinger and Long (2005) provide evidence for college instruction.

- the health-oriented profile Nature & Health (NH)
- the social science-oriented profile Economics & Society (ES)
- the humanities-oriented profile Culture & Society (CS)

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

Table 1 shows the subjects that differ across the different study profiles and the number of teaching hours assigned to these subjects in the last three years of secondary school. Mathematics is the only subject taught at different levels in each track, where D is the most advanced math course followed by B, A and C. The order of math difficulty is therefore NT>NH>ES>CS. Panel A of Table 2 shows the strong correlation between the study profile in secondary school and the choice of major in tertiary education. Most NT graduates 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 in law, and most CS graduates choose a subject in the humanities, social sciences or law.[17]

Generally, NT is viewed as the most challenging and prestigious study profile, followed by NH and ES, and CS as the least demanding and prestigious study profile. This is related to the difficulty and amount of mathematics and science in the curriculum. In other countries in which students can choose study profiles in school, the prestigiousness of the profiles is also often highly correlated with their math intensity (see for example Pautler (1981) for France).

The prestigiousness of study profiles is related to the likelihood of going to university, see panel B of Table 2. The prestige of study profiles is also related to the academic performance of students selecting the various profiles.

The top panel of Table 3 shows mean values of measures of students' ability by study profile. This information is based on the students in our sample. According to all five of our ability measures, the students who choose NT have higher ability than the students who choose NH, who in turn have higher ability measures than the students who choose ES. Students who choose CS score lowest on four of the five measures.

This ordering of profiles is also reflected in the opinions of the students from our study. In a short questionnaire we asked the students to rank the four study profiles by asking "Which profile do the best students pick?". The bottom panel of Table 3 shows that their responses concur with the general opinion. A majority of over 70 percent of students believes NT is chosen by the best students. A majority of students ranks the NH study profile second and ES third. More than 80 percent of students rank CS as the profile most chosen by the weakest students. The rankings of boys and girls are very similar. We also asked students to rank the four study profiles in terms of future earnings. The picture that emerges is also very similar. The exact question was "With which profile do you think you would earn most in ten year's time? Rank the profiles from 1 to 4 where 1 means that you would earn most if you chose that profile and 4 that you would earn least if you

[17]Some studies actually restrict entry to certain profiles or courses within profiles. For example, medical schools require NT or NH; to study Math, having taken at least Math B in high school is required.

Table 1. Subjects and teaching hours per study profile

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

Note: The table lists the subjects per profile and the number of teaching hours per subject during the last three years of the pre-university track. In addition all students take the following non-profile 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 profile specific subjects and half on common subjects. Source: Ministry of Education, Culture and Science.

Table 2. Study profiles by tertiary education choices and gender (percentages)

		NT	NH	ES	CS
A	**Undergraduate major**				
	Humanities	9	6	8	30
	Social Sciences	2	9	19	34
	Law	1	4	20	20
	Economics and Business	15	8	46	5
	Science and Engineering	64	18	2	0
	Health Care	7	48	1	1
	Other	2	7	4	9
B	Going to university	81	72	69	60
C	Boys	35	21	38	6
	Girls	10	34	32	24

Source: Statistics Netherlands (CBS). The data from the top rows are from 2006. The data from the bottom rows are from 2009, where we exclude choices of combined profiles.

Table 3. Descriptive statistics about profiles

By chosen profile	NT	NH	ES	CS	Difference
GPA (1-10)	7.12	7.09	6.65	6.61	0.00
Math grades (1-10)	7.25	6.73	6.20	6.21	0.00
Relative math (0-1)	0.23	0.35	0.46	0.49	0.00
Math difficulty (0-10)	1.95	3.62	4.90	5.30	0.00
Math quartile (1(best)-4)	1.52	1.98	2.50	2.67	0.00
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	

Note: Top rows: Average ranking of study profiles, and in parentheses, the fraction of students who rank that profile first (for NT), second, third or fourth for NH, ES and CS, respectively. Bottom rows: Average characteristics of subjects who chose that profile. Grades are out of 10 with higher numbers being better grades. Math difficulty goes from 0 - very easy to 10 - very hard. Math quartile goes from 1 - best 25% to 4 - worst 25%. The last column reports p-values from Kruskal Wallis tests.

chose that profile." This question was only asked to students in two of the four schools and the percentages are therefore based on 181 observations. 50 percent think that NT gives the best salary prospects, 27 percent think NH, 20 percent ES and 2 percent CS.

In the remainder of the paper, we order the profiles in terms of decreasing prestigiousness: NT $>$ NH $>$ ES $>$ CS. As a robustness check, we use in the appendix for each student the ranking they gave to their chosen profile in terms of which profiles the best students choose. That is, if a student ranked, say, CS as the profile chosen by the best students, followed by ES, NH and NT (so, the reverse order) and chose profile CS for herself, we categorize that student as choosing the most "own prestigious" profile (rank 4). If this student chose ES, we would rank her choice as 3 and so on. The results remain qualitatively the same.

Panel C of Table 2 shows that in the Netherlands, profile choices differ markedly between the sexes. Boys are more likely to choose more prestigious study profiles. Compared to girls, boys are more than three times as likely to choose the most prestigious profile, NT, and only a fourth as likely to choose the least prestigious profile, CS. The fact that girls are disproportionally more likely to choose CS has prompted a debate with the minister for education even proposing to eliminate the profile altogether. This idea was ultimately rejected and the profiles remain as they are for now.[18]

3.2 Academic Data

Before we assess the profile choices of boys and girls in our data, we present the academic data we received from schools.

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

Ability. The first three rows of Panel A in Table 4 show that girls have significantly higher GPA than boys, while there is no significant gender difference in the absolute or relative grade for mathematics. The last two rows of Panel A in Table 4 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.

Profile choice. We have two sources of information about students' profile choices. In the questionnaire we asked students which profile they expected to choose. The schools 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 profiles. 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 profile, students take Mathematics B, albeit only at 520 hours. Furthermore, Physics is not required. In the ES/CS profile, students replace one of the CS-electives with the economics course. As such, the combined profiles are somewhat in between the pure profiles, though a little closer to NT and ES, respectively. For the main analysis of this paper we use for the students in combined profiles, the chosen profile as stated in the questionnaire.[19] However, since one can argue that the NT/NH profile is closer to NT, and the ES/CS closer to ES, we reestimate all regressions using this alternative definition of profile choice in the appendix. As a further robustness check, in the appendix, we show results where we treat NT/NH and ES/CS as separate categories.[20] The results remain qualitatively the same in both specifications.

While academically boys and girls are very comparable, girls make significantly different profile choices from boys. The lower part of Table 4 shows profile choices by gender in our sample of 362 students. The pattern is similar to the pattern observed in the national statistics (see Table 2). The NT profile is much more popular among boys than girls, while the opposite holds for NH. The ES profile is slightly more popular among boys than girls, and girls are much more likely than boys to choose the least prestigious profile, CS. Note that in our sample, boys and girls are as likely to choose one of the science profiles compared to one of the society profiles. This is not the case in the national statistics, where girls are overall more likely to choose a society profile, while boys are overall more likely to choose a science profile.[21]

[19] 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 profile to place students that chose a combination profile into "pure" profiles.

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

[21] There are two potential reasons for the disparity in gender differences in profile choices between our set of four schools and the national statistics. First, students at schools in and around Amsterdam may differ from the average Dutch student. Second, the national statistics relate to cohorts that will graduate in 2009, while students in our sample will graduate in 2014. In addition, there seems to be a general trend away from CS.

Table 4. Descriptive statistics by gender

	Scale	Boys	Girls	p-value
A: Ability				
GPA	1(lowest) - 10(highest)	6.80	6.97	0.01
Math grade	1(lowest) - 10(highest)	6.67	6.59	0.49
Math relative	0-1	0.38	0.37	0.88
Math quartile (1-4)	1(best) - 4(worst)	1.97	2.25	0.03
Math difficulty (0-10)	0(very easy) - 10(very hard)	3.41	4.18	0.01
B: Profile choices				
Nature & Technology (NT)	dummy	0.40	0.17	
Nature & Health (NH)	dummy	0.12	0.36	
Economics & Society (ES)	dummy	0.39	0.32	
Culture & Society (CS)	dummy	0.08	0.15	0.00
Number of observations		177	185	

Note: The last column reports p-values from t-tests for continuous variables and from a Fisher's exact test for categorical variables.

3.3 Gender Differences in Prestige of Chosen Profiles

To more precisely understand gender differences in the prestige of the chosen study profiles, we show in Table 5 ordered probit regressions where we order profiles from most to least prestigious: NT>NH>ES>CS. In this and all following analyses, we standardize all non-binary control variables to make the coefficients comparable. Table 10 in the online appendix provides the mean and standard deviations of all our control variables. The first column shows that boys are significantly more likely than girls to choose a prestigious profile. Being female bridges almost 20 percent of the distance between the most and the least prestigious profiles (this is shown in the penultimate row by F/(C3-C1), the female coefficient divided by the distance between the first and the third ordered probit cutoffs). Inclusion of objective ability variables (column (2)) increases the gender gap to almost 22 percent of the distance between the least and the most prestigious profiles. Note that the coefficient on female is larger (in absolute value) than on the GPA. An increase of one standard deviation in GPA corresponds to bridging 11 percent of the gap between the most and the least prestigious profile.

When we add students' perceptions about their mathematics ability in column (3), the gender gap shrinks but remains large and highly significant. While 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 significant gender difference in study profile choice, with girls choosing less prestigious profiles than boys.[22]

Table 11 in the appendix shows that the results are very similar when we classify an NT/NH

[22] 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 -0.296 (s.e. 0.105, $p<0.01$) without any controls. The magnitude of the effect increases to -0.337 (s.e. 0.101, $p< 0.01$) when we add the controls from column (2) in Table 5, which is larger than the coefficient on standardized GPA which is 0.207 (s.e. 0.059, $p<0.01$). When we add all the controls from column (3) the gender coefficient is -0.225 (s.e. 0.095, $p<0.05$), again larger than the coefficient on the GPA of 0.175 (s.e. 0.058, $p< 0.01$).

Table 5. Gender and profile choice

	Ordered probit (NT>NH>ES>CS)			NT vs. Rest	Rest vs. CS
	(1)	(2)	(3)	(4)	(5)
Female	-0.342***	-0.433***	-0.315**	-0.207***	-0.061*
	(0.114)	(0.124)	(0.125)	(0.047)	(0.033)
Math Grade		0.189	0.001	0.038	-0.051*
		(0.149)	(0.158)	(0.068)	(0.030)
GPA		0.218**	0.201**	0.027	0.035*
		(0.092)	(0.092)	(0.041)	(0.018)
Rel. Math Gr.		-0.152	-0.106	-0.062	-0.036
		(0.126)	(0.128)	(0.063)	(0.025)
Math Difficulty			-0.220**	-0.110**	-0.021
			(0.087)	(0.046)	(0.017)
Math Quartile			-0.321***	-0.154***	-0.029*
			(0.072)	(0.038)	(0.015)
Cut 1 (C1)	-1.367***	1.806	-0.535		
Cut 2 (C2)	-0.251***	3.038***	0.794		
Cut 3 (C3)	0.404***	3.787***	1.610		
Female/(C3-C1)	-0.193***	-0.219***	-0.147***		
N	362	362	362	362	362

Note: Dependent variable in columns (1) to (3): Profile choice, where NT>NH>ES>CS. Coefficients from ordered probit regressions. Dependent variable in column (4): dummy variable NT=1. Dependent variable in column (5): dummy variable "not CS"=1. Marginal effects in columns (4) and (5) from probit regressions. Robust standard errors in parentheses; *, ** and *** denote significance at 10, 5 and 1 percent, respectively. The margins are taken for a male student and mean values of the other variables.

combined choice as NT, and an ES/CS choice as ES, instead of using the students' answer in the questionnaire to attribute combined profile choices to one of the four baseline study profiles. The results are also robust to treating the combined profiles as their own category, where combined profiles are ordered between the baseline study profiles, that is, NT>NT/NH>NH>ES>ES/CS>CS. Finally, 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 11 in the appendix).

To provide additional insights on the magnitude of the gender difference in profile choice, we run probit regression on choosing the most prestigious profile, NT, compared to any other profile, controlling for objective and subjective academic performance. Column (4) in Table 5 shows that girls are 21 percentage points less likely to choose NT, a significant difference. When we redo the exercise for choosing the least prestigious profile, CS, compared to any other profile, the marginal coefficient shows that female students are 6 percentage points more likely to choose CS than boys, again a significant difference (column (5)).

4 Gender Differences in Competitiveness

In this section we analyze gender differences in competitiveness as well as confidence and risk aversion.

4.1 Experimental Data

Competitiveness. Panels A and B of Table 6 report mean values of performance in Rounds 1 and 2 and of tournament entry, separately for boys and girls. In Round 1 boys perform significantly better than girls. In the second round when students' payment is based on the tournament, there is no significant difference in performance. Since students compete only against students in their own class, we compute for each student the chance to win the tournament in Round 2 given their performance and that of their classmates.[23] The average chance to win the tournament is slightly but not significantly higher for boys than for girls. Provided the performance in Round 3 is not lower than that in Round 2, then every student with a chance of winning the tournament of 25 percent and higher has higher expected earnings when choosing to enter the tournament in Round 3. This would result in 38 percent of the boys and 35 percent of the girls entering the tournament, an insignificant difference. Actual tournament entry shows a very different pattern; we find that 49 percent of boys and less than half as many, 23 percent of girls, enter the tournament. This difference is significant (and significantly different from optimal entry p=0.01).[24]

[23] 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 one thousand different comparison groups of three from a participants' own class. If two performances were tied for first place, a 0.5 win was assigned (1/3 in case of three tied performances and 0.25 in case of four).

[24] In Round 3, subjects who compete solve on average 9.75 correct sums while those who do not compete solve 7.92 (p=0.00). 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=0.25 and p=0.65, respectively).

Table 6. Descriptive statistics by gender

	Scale	Boys	Girls	p-value
A: Performance				
Performance Round 1 (piece rate)	number of correct answers	6.60	5.94	0.03
Performance Round 2 (tournament)	number of correct answers	7.90	7.42	0.15
Chance of winning Round 2 (tournament)	[0,1]	0.27	0.24	0.24
B: Competitiveness				
Actual tournament entry	dummy	0.49	0.23	0.00
Optimal tournament entry	dummy	0.38	0.35	0.59
C: Confidence				
Actual guessed rank	1(best) - 4 (worst)	2.14	2.56	0.00
Optimal guessed rank	1(best) - 4 (worst)	2.39	2.55	0.24
Guesses to be the best	dummy	0.32	0.11	0.00
Optimal to guess to be the best	dummy	0.25	0.22	0.46
D: Risk Attitudes				
Lottery choice	1(no risk) - 5(highest risk)	3.46	2.99	0.00
Risk taking	1(avoid risk) - 10(seek risk)	6.52	5.96	0.00
Number of observations		177	185	

Note: The last column reports p-values from t-tests for continuous variables and from a Fisher's exact test for categorical variables.

Confidence. Panel C of Table 6 reports that the average guessed rank is 2.14 for boys and 2.56 for girls, with the two distributions being significantly different. We find that 32 percent of the boys and 11 percent 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 the performances of the other students in their class.[25] Using the optimal guessed rank, there would be no significant gender difference in overall beliefs or the guess to be the best. An ordered probit regression of the guessed rank on the optimal guessed rank and a female dummy delivers a female coefficient of 0.496 (s.e. 0.117, $p = 0.00$).[26] This confirms that girls, given their relative performance, are significantly less confident about their relative performance than boys.

Risk attitudes. Panel D in Table 6 shows that boys on average choose a significantly more risky lottery. On the general risk tolerance question boys also score on average significantly higher. The correlation between the two risk measures is 0.42 in the whole sample ($p<0.01$), and 0.45 and 0.34 in the sub-samples of boys and girls, respectively ($p<0.01$ in both cases).

[25] We compute the optimal guessed rank through simulation. We randomly draw a thousand 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.

[26] The coefficient on the optimal guessed rank is 0.653 (s.e. 0.060, p=0.00).

Table 7. Determinants of tournament entry

	(1)	(2)	(3)	(4)	(5)
Female	-0.261***	-0.191***	-0.166***	-0.153***	-0.139**
	(0.051)	(0.055)	(0.055)	(0.056)	(0.058)
Tournament	0.054***	0.022	0.020	0.024	0.015
	(0.020)	(0.021)	(0.021)	(0.021)	(0.021)
T - PR	-0.028**	-0.022*	-0.018	-0.018	-0.014
	(0.013)	(0.013)	(0.014)	(0.014)	(0.014)
Win Prob	0.231	0.062	0.041	-0.006	0.080
	(0.204)	(0.210)	(0.211)	(0.215)	(0.226)
Guessed rank		-0.275***	-0.276***	-0.258***	-0.253***
		(0.041)	(0.041)	(0.041)	(0.041)
Lottery			0.098***	0.043	0.048
			(0.031)	(0.032)	(0.033)
Risk-taking				0.156***	0.167***
				(0.035)	(0.036)
Math grade					0.165**
					(0.075)
GPA					-0.089*
					(0.048)
Math Relative					0.027
					(0.066)
Math quartile					0.068
					(0.042)
Math difficulty					-0.025
					(0.045)
N	362	362	362	362	362

Dependent variable: Round 3 choice of compensation scheme (1-tournament and 0-piece rate). The table presents marginal effects of coefficients of a probit regression evaluated at a male student with a 0.25 chance of winning (the rest of the variables are evaluated at the sample mean). Standard errors of the marginal coefficients are in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ of the underlying coefficient.

4.2 Gender differences in competitiveness

To assess gender differences in competitiveness, Table 7 shows marginal effects from probit regressions of tournament entry in Round 3. Girls have a 26 percentage point lower probability of entering the tournament compared to boys, when only controlling for performance in Round 1, the difference in performance between Rounds 1 and 2 and the chance of winning in Round 2 (column (1)). This is very much in line with Niederle and Vesterlund (2007) and the large resulting literature (see Niederle and Vesterlund, 2011).

Adding the guessed rank as a measure of confidence to the probit regression on tournament entry, column (2) shows that the gender effect drops to 19 percentage points, a still highly significant difference.[27] Adding the lottery choice variable slightly reduces the gender gap in tournament

[27] 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

entry by an additional 2 percentage points to 17 percentage points (compare columns (2) and (3)). Adding the questionnaire-based risk measure reduces the gender gap by another 2 percentage points (compare columns (3) and (4)). Finally, when we also include measures of academic performance and of perceived mathematical ability, the gender gap in tournament entry is once more slightly reduced, leaving a significant gender gap in tournament entry of 14 percentage points (column (5)).

In summary, the students in our sample follow the standard gender differences in choice of competition (see Niederle and Vesterlund, 2011). Controlling for performance, girls are about 26 percentage points less likely to enter the tournament. Boys have significantly more optimistic views about their relative performance than girls, and these gender differences in confidence account for slightly more than a quarter of the gender gap in tournament entry. Risk attitudes, whether measured by a lottery choice or a simple questionnaire item, while significantly predicting tournament entry, reduce the gender gap in competitiveness only by a much smaller amount once we control for confidence.

5 Can competitiveness account for gender differences in prestige of chosen profiles?

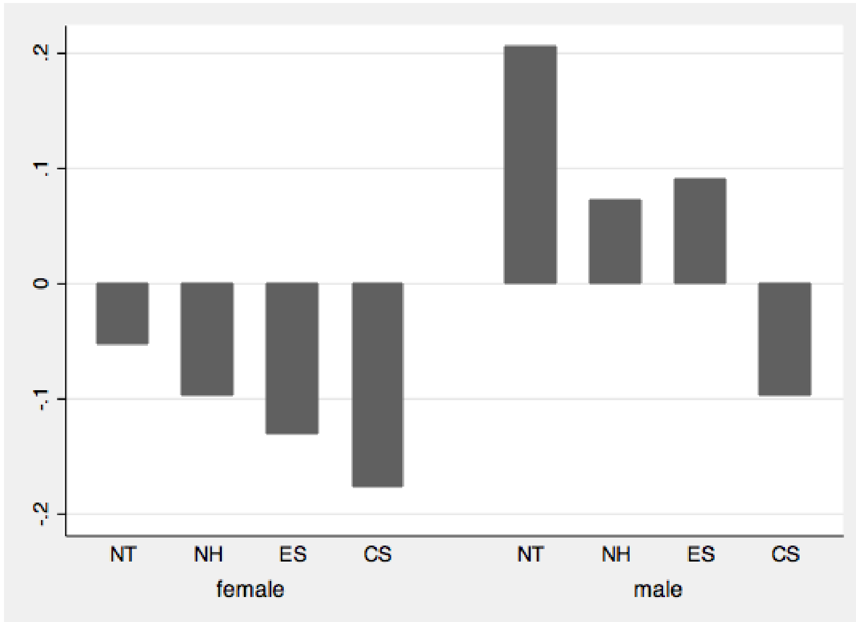
The students in our sample represent a classical situation. Boys and girls agree on which academic profiles are the most prestigious, where prestigiousness perfectly correlates with math intensity. They do not differ in math grades, and if anything, girls have a slightly higher GPA than boys. Despite these facts, girls are significantly less likely to choose the prestigious NT profile, and, in turn, significantly more likely to choose the least prestigious profile, CS. Gender differences in study profile choices remain significant when controlling for academic performance and perceived mathematical ability. The students in our sample also exhibit the standard gender gap in competitiveness.

In this section we assess whether competitiveness correlates with the prestigiousness of the study profile choice and, more importantly, whether gender differences in competitiveness can help account for the gender gap in chosen profiles. We also confirm that our results are robust to the inclusion of different sets of ability control variables, as well as controls for confidence and risk attitudes which we showed to be correlated with tournament entry.

As a preliminary analysis, Figure 1 shows for each profile the mean competitiveness of boys and girls who chose that profile. In the figure, competitiveness is measured as the residual from a regression of tournament entry on performance in Rounds 1 and 2 of the experiment and the chance of winning the Round 2 tournament. The figure shows that within each sex, students who enter the NT track are the most competitive students, followed by NH and ES, while those who enter the CS track are the least competitive ones. The ranking is more pronounced among boys than among girls. The figure suggests that competitiveness as measured by the short in-class experiment correlates with study profile choice and as such may help account for the gender difference in those choices.

4 percent and a gap of 25 percentage points remains. The coefficient on female is -0.251, (s.e. 0.051, $p<0.01$), not very different from the -0.261 from column (1). Adding all measures on 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 19.4 (s.e. 5.5, $p<0.01$) percentage points less likely to enter the tournament.

Figure 1. Tournament entry by gender and subsequent profile choice (conditional on performance)



To investigate the impact of competitiveness on gender differences in profile choice in an intuitive way, we compare the impact of gender on educational choices for different subpopulations split by gender and tournament entry. If the gender gap in profile choice is unrelated to competitiveness, the impact of gender on profile choice should, for example, be the same for the subsample made up of competitive boys (Comp Boys) and non-competitive girls (N-comp Girls) as for the subsample made up of non-competitive boys and competitive girls.

This idea is explored in Table 8 which reports coefficients of regressions of profile choice on a female dummy (and controls for performance in the experiment) for the various subsamples. The top part reports ordered probit estimations that rank profiles by prestige: NT>NH>ES>CS. The table shows that the gender gap in profile choice, which is significant for the whole sample, varies strongly with competitiveness as measured by tournament entry. The gender gap in profile choice increases with the competitiveness of boys and decreases with the competitiveness of girls. When we consider competitive boys and non-competitive girls, gender bridges about 36% of the gap between choosing the most and the least prestigious profile (see column (2) of the upper half of Table 8). When, on the other hand, we consider boys who chose the piece rate and girls who entered the tournament, there is no significant gender difference in profile choices. Furthermore, the change in the gender dummy between the group of competitive boys and non-competitive girls and the other way round is significant at the 1%-level.

The probit models in the lower part of Table 8 give a more detailed view on this result. We first assess the probability of choosing the most prestigious NT profile compared to any other study profile. When we consider only competitive boys and non-competitive girls, girls are 34 percentage points less likely to choose NT. When instead we consider competitive girls and non-competitive boys, girls

Table 8. Gender effects by subsample

	(1) Ordered probit	(2) Female/(C3-C1)	N
(1) Comp B & n-comp. G	-0.67***	-0.36***	230
	(0.16)		
(2) Comp. B & comp. G	-0.49**	-0.28***	129
	(0.20)		
(3) N-comp. B & n-comp. G	-0.16	-0.09	233
	(0.15)		
(4) N-comp. B & comp. G	0.03	0.02	132
	(0.20)		
(5) Whole sample	-0.35***	-0.20***	262
	(0.12)		
P-value (1) vs (4)	0.00	0.01	

Probit (marginal effects)	(1) NT vs Rest	(2) N vs S	(3) Rest vs CS	(4) Best vs Rest
(1) Comp B & n-comp. G	-0.34***	-0.10	-0.13***	-0.31***
	(0.07)	(0.07)	(0.04)	(0.07)
(2) Comp. B & comp. G	-0.30***	-0.00	-0.09*	-0.23**
	(0.08)	(0.09)	(0.05)	(0.09)
(3) N-comp. B & n-comp. G	-0.17***	0.04	-0.02	-0.15**
	(0.06)	(0.07)	(0.05)	(0.06)
(4) N-comp. B & comp. G	-0.09	0.13	0.01	-0.08
	(0.09)	(0.10)	(0.06)	(0.09)
(5) Whole sample	-0.24***	-0.00	-0.07**	-0.21***
	(0.05)	(0.05)	(0.03)	(0.05)
P-value (1) vs (4)	0.01	0.04	0.05	0.02

Note: Coefficients are from regressions of profile choice on a female dummy and controls for performance in Rounds 1 and 2 of the experiment and the chance of winning the Round 2 tournament; robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; p-values are bootstrapped. For the ordered probit regression, F/(C3-C1) represents the relative size of the female coefficient to the cuts provided by the ordered probit regression between choosing least and second to least most prestigious profile and choosing the second to most prestigious profile, respectively. For probit regressions, the value is 1 for the left variable (e.g. NT in column (1)) and 0 otherwise. The margins are taken at mean values of the other variables.

are only 9 percentage points less likely to choose the NT profile, a difference that is not significant. Furthermore, the gender difference in choices is significantly smaller when we consider competitive girls and non-competitive boys than when we consider competitive boys and non-competitive girls. The results are qualitatively similar when we either consider choices between the Nature and the Society profiles (the top two versus the bottom two in terms of prestige), or when we consider the option to choose CS, the least prestigious profile, compared to any other profile. Finally, we consider the student specific ordering of study profiles. Specifically, for each student we ask whether they pick the profile that they deem to be the one chosen by the best students, or another profile. In all cases the results are very similar. Gender differences are reduced when we reduce the competitiveness of boys and increase the competitiveness of girls.

These results show that gender differences in profile choice are strongly related to gender difference in competitiveness. We now turn to detailed regressions to confirm that the impact of competitiveness on profile choice is robust when controlling for academic variables.

Table 9 shows ordered probit regressions on the ranked profile choice. All columns include controls for performance in Rounds 1 and 2 of the experiment and the chance of winning the Round 2 tournament. Column (1) shows that being female bridges 19.5 percent of the gap between choosing the least and the most prestigious profile, controlling for performance in the experimental task only. In column (2) we add competitiveness which takes a value of 1 for students who entered the tournament in Round 3 of the experiment and 0 if the student chose the piece rate. The first result is that the coefficient on competitiveness is significant, with more competitive students choosing more prestigious profiles. The decision to enter the tournament bridges 19.4 percent of the gap between choosing the least and the most prestigious profiles, more than the gender of the student. In fact, when there is no other information available, knowing a students' competitiveness is a slightly better predictor of study profile choice than knowing their gender.[28]

The main finding is that adding competitiveness in column (2) significantly changes the effect of being female from 0.195 to 0.150, a reduction of 23 percent.[29] This shows the importance of competitiveness to predict study profile choices, as well as account for the gender gap in those choices.

Over the following columns of Table 9, we add controls for objective and perceived academic ability, separately and jointly. The coefficient of tournament entry is robust and stays significant throughout, with more competitive students selecting more prestigious profiles. The coefficient on competitiveness is under all specifications comparable to the one on gender, ranging from 75 to 130 percent of the gender coefficient. This confirms the relevance of competitiveness to account for educational choices.[30]

[28] When running the ordered probit on tournament entry only (controlling for performance), Entry/(C3-C1) is 0.237 compared to Female/(C3-C1) in column (1) of Table 9 which is 0.195.. Similarly, using OLS to regress ordered profile choice on a female dummy only yields an R^2 of 2.16 percent whereas regressing on the entry dummy yields an R^2 of 2.93 percent. Adding entry on top of female raises theR^2 from 2.16 percent to 4.01 percent.

[29] We use bootstrap to calculate the significance of this difference. This is done by resampling 10,000 times with replacement, keeping the number of male and female subjects constant. We then count the fraction of differences between column (2) and column (1)'s Female/(C3-C1) variable that are negative or zero.

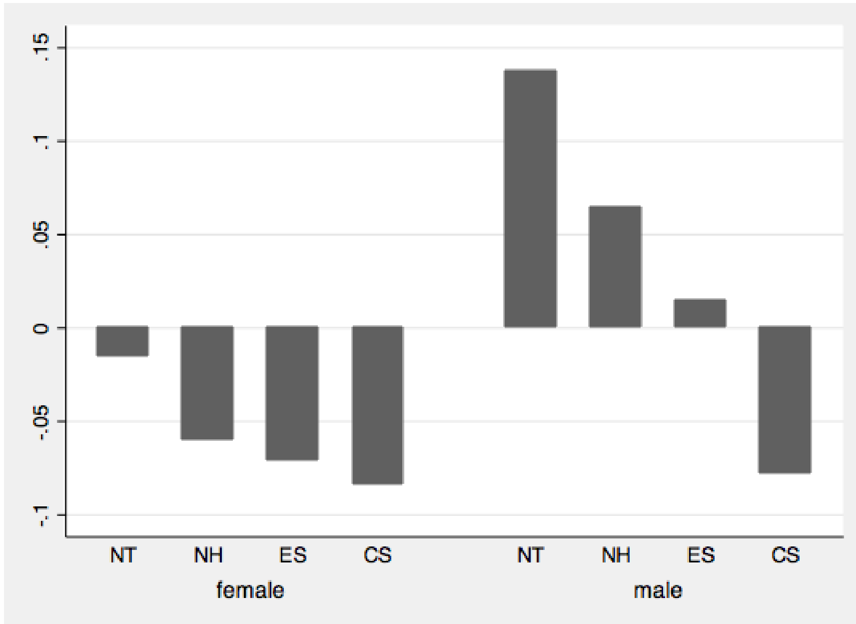
[30] When running the ordered probit on tournament entry and all the controls (except the female dummy) in column

Table 9. Profile choice: ordered probit regression

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Female	-0.350***	-0.272**	-0.462***	-0.406***	-0.219*	-0.157	-0.343***	-0.284**	-0.359***	-0.308**	-0.305**	-0.254*	-0.319**	-0.276**
	(0.116)	(0.120)	(0.127)	(0.131)	(0.119)	(0.126)	(0.129)	(0.134)	(0.130)	(0.133)	(0.131)	(0.134)	(0.131)	(0.133)
Entry		0.353***		0.300**		0.295**		0.315**		0.400***		0.331**		0.404***
		(0.130)		(0.131)		(0.132)		(0.134)		(0.144)		(0.142)		(0.152)
Math grade			0.185	0.154			-0.007	-0.040	-0.002	-0.038	0.012	-0.027	0.016	-0.025
			(0.149)	(0.150)			(0.158)	(0.157)	(0.158)	(0.157)	(0.159)	(0.157)	(0.159)	(0.157)
GPA			0.227**	0.252***			0.213**	0.239***	0.212**	0.243***	0.201**	0.223**	0.201**	0.228**
			(0.092)	(0.092)			(0.091)	(0.091)	(0.092)	(0.091)	(0.092)	(0.092)	(0.093)	(0.092)
Math relative			-0.156	-0.156			-0.115	-0.116	-0.119	-0.125	-0.098	-0.097	-0.102	-0.106
			(0.128)	(0.128)			(0.131)	(0.131)	(0.131)	(0.130)	(0.133)	(0.132)	(0.132)	(0.131)
Math difficulty					-0.333***	-0.321***	-0.209**	-0.203**	-0.211**	-0.207**	-0.220**	-0.220**	-0.220**	-0.222**
					(0.078)	(0.078)	(0.087)	(0.086)	(0.088)	(0.088)	(0.089)	(0.088)	(0.090)	(0.090)
Math quartile					-0.356***	-0.361***	-0.329***	-0.335***	-0.334***	-0.350***	-0.335***	-0.341***	-0.340***	-0.355***
					(0.076)	(0.077)	(0.074)	(0.075)	(0.075)	(0.076)	(0.075)	(0.075)	(0.075)	(0.076)
Guessed rank									0.054	0.132			0.053	0.121
									(0.078)	(0.084)			(0.081)	(0.086)
Risk											-0.057	-0.100	-0.048	-0.090
											(0.068)	(0.069)	(0.068)	(0.069)
Lottery											0.146**	0.137*	0.145**	0.133*
											(0.072)	(0.072)	(0.072)	(0.072)
Cut 1	-1.048***	-1.003***	1.943*	2.065*	-2.535***	-2.488***	-0.337	-0.220	-0.165	0.240	-0.190	-0.312	0.022	0.149
Cut 2	0.080	0.138	3.190***	3.323***	-1.229***	-1.172***	1.007	1.137	1.179	1.599	1.167	1.057	1.379	1.520
Cut 3	0.747***	0.812***	3.953***	4.091***	-0.425*	-0.362	1.837	1.974	2.010	2.441*	2.001	1.899	2.214*	2.366*
Female/(C3-C1)	-0.195***	-0.150**	-0.230***	-0.200***	-0.104**	-0.074	-0.158***	-0.130**	-0.165***	-0.140**	-0.139**	-0.115**	-0.146***	-0.124**
Diff.	23.1%		12.9%		29.0%		17.8%		15.3%		17.4%		14.7%	
Bootstrap p-value	0.003		0.013		0.014		0.012		0.006		0.014		0.012	
Observations	362	362	362	362	362	362	362	362	362	362	362	362	362	362

Note: Coefficients are from ordered probit regressions, where NT>NH>ES>CS. All specifications include controls for performance in rounds 1 and 2 and the chance of winning. Coefficients are from ordered probit regressions; robust standard errors in parentheses; p-values for F/(C3-C1) and Dif. are bootstrapped; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The p-values for the impacts of confidence (comparing columns (7) and (9)) and risk attitudes (comparing columns (7) and (11)) on the gender gap (Female/(C3-C1)) are 0.74 and 0.06 respectively.

Figure 2. Tournament entry by gender and subsequent profile choice (conditional on performance, confidence and risk attitudes)



Pairwise comparisons between Columns (3) and (4), (5) and (6), and (7) and (8) confirm that competitiveness explains a substantial part of the gender gap in profile choice; 13 percent when controlling for academic performance (columns (3) and (4)), 29 percent when controlling for believed mathematical ability (columns (5) and (6)) and 18 percent when controlling for both subjective and objective academic performance (columns (7) and (8)). These changes in the gender coefficient upon inclusion of our competitiveness measure are significant for all specifications. This confirms that competitiveness can account for a substantial part of the gender gap in profile choice even after controlling for academic performance and perceived mathematical ability.

The online appendix shows that the 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.

We have previously shown that tournament entry is partially explained by confidence and risk attitudes. These attributes could also conceivably be correlated with profile choice. In what follows, we assess the impact of competitiveness on profile choice and its impact on gender differences when accounting for confidence and risk attitudes. For a preliminary analysis, Figure 2 shows for each profile the mean competitiveness of boys and girls who chose that profile. But here, competitiveness is measured as the residual from a regression of tournament entry not only on the performance measures, but also the guessed rank and the risk measures. The ranking of profiles in terms of

(7) of Table 9, Entry/(C3-C1) is 0.248 compared to Female/(C3-C1) in column (7) which is 0.158. Adding a female dummy to an OLS regression with all the controls used in Column (7) increases the R^2 by 1.25 percentage points while adding entry raises it by 1.43 percentage points (adding entry on top of female raises the R^2 by a further 0.94 percentage points).

competitiveness stays intact which indicates that the impact of competitiveness on the profile choice is not due to the impact of risk attitudes and confidence.

In columns (9) to (14) of Table 9 we add controls for confidence and risk attitudes to ordered probit regressions on profile choice ranked by prestigiousness. The main result is that the coefficient of tournament entry and its effect on the gender gap in profile choice remains robust and stays significant throughout. Pairwise comparisons between columns (9) and (10), (11) and (12), and (13) and (14) confirm that competitiveness explains a substantial part of the gender gap in profile choice, when controlling for either confidence, risk attitudes, or both. Column (8) shows that competitiveness reduces the gender gap in prestigiousness of profile choice by 18 percent when controlling for actual and perceived academic ability. When we control in addition for both confidence and risk aversion measures, competitiveness still reduces the gender gap by 15 percent (column (13) versus column (14)). When we add all three behavioral measures, the gender gap in profile choice is reduced by 22 percent (column (7) versus (14)). Using only competitiveness resulted in a reduction of the gender gap that is 82 percent of the size of the effect of all psychological attributes. Risk aversion and confidence jointly reduce the gender gap by 35 percent of the effect of all psychological attributes (column (7) versus column (13)). These results imply that the reduction in the gender gap when controlling for competitiveness is not due to an impact of confidence or risk attitudes. The result holds when we consider other specifications of combination profiles or use the students' own ranking of prestigiousness of profiles (see appendix).[31]

Finally, we can consider the effect of confidence and risk measures on profile choice and its gender gap separately. Column (9) shows that confidence (as measured by the guessed rank in the Round 2 tournament, while we keep controls for the performance in Rounds 1 and 2) has no significant influence on the prestige of the chosen profile. Comparing columns (7) and (9) reveals that the inclusion of the confidence measure has also no impact on the gender gap in choices (which in fact increases slightly). These conclusions are mirrored when we control in addition for competitiveness (see columns (8) and (10)). The results are robust to other specifications of profiles.

Column (11) of Table 9 shows that risk attitudes correlate with the prestige of the chosen profile. Students who opted for a more risky lottery enroll in more prestigious study profiles. Comparing columns (7) and (11) shows that adding risk attitudes reduces the gender gap by around 12% (this reduction is significant at the 10 percent level). The effects of competitiveness and risk attitudes on the gender gap in profile choice are almost orthogonal; adding only competitiveness reduces the gender gap by 17.8 percent (compare columns (7) and (8)); adding only risk attitudes reduces the gender gap by 12 percent (compare columns (7) and (11)); adding competitiveness and risk attitudes together reduces the gender gap by 27.2 percent (compare columns (7) and (12)). This effect is somewhat weaker (and not always significant) in our alternative specifications where it ranges from 6 to 10 percent (see online appendix).

[31] When we classify combination profiles as the more prestigious profile in the combination, competitiveness alone generates in the gender gap in study profile choices that is 66 percent of the overall size of the reduction generated by all three psychological attributes. The corresponding number is 75 percent and 100 percent when we either classify combination profiles as separate choices or when we use for each student their own pranking of prestigiousness, respectively.

6 Conclusion

This study examines whether experimentally measured gender differences in competitiveness can account for gender differences in career choices. We analyze the first important career choice of young people in the Netherlands, for which we observe substantial gender differences. At the end of 9^{th} grade, students in the pre-university track choose between four study profile which are ranked according to difficulty, prestige and math intensity in the following order: a science profile (NT), a health profile (NH), a social science profile (ES) and a humanities profile (CS). In our sample of all such students from four schools in and around Amsterdam, 40 percent of the boys (in our data) choose the challenging NT profile, while only 17 percent of the girls do so, and while 15 percent of the girls choose CS, the least prestigious profile, only 8 percent of the boys do so. Ordered probit regressions confirm that girls are significantly less likely to select a prestigious study profile, despite the fact that girls are as good at math as the boys, and actually have a higher GPA.

We assess the competitiveness of students through a classroom experiment several months before they make their profile choice and while they still share the same classroom experience in school. We use the Niederle and Vesterlund (2007) design, where students, after performing in a simple addition task under a piece rate and a tournament scheme, can select the payment scheme for their final round. We find significant gender differences in competitiveness (defined as tournament entry after controlling for performance in the experiment).

We find that competitiveness varies strongly and significantly across study profiles, with students that are more competitive selecting more prestigious profiles. Ordered probit regressions confirm that competitiveness predicts the prestigiousness of the chosen profile, even when we control for academic performance and perceived mathematical ability. The main result is that our simple measure of competitiveness can account for 18 percent of the gender gap in profile choice after we control for objective and subjective academic performance. When we also include controls for confidence and risk attitudes we still find that inclusion of competitiveness reduces the gender gap in profile choices by 15 percent. To evaluate the size of these effects note that competitiveness is comparable to (and varying between 75 to 130 percent of) the effect of gender on profile choices. These results are a direct demonstration of the field relevance of laboratory findings on gender differences in competitiveness. More such demonstrations should follow so that labor economists take laboratory findings on gender differences in psychological attributes into account in their study of gender gaps (cf. Bertrand, 2011).

This paper is part of a small but growing literature that aims to predict economic outcomes outside of the laboratory with laboratory measures, see e.g. Karlan (2005), Ashraf et al. (2006), Meier and Sprenger (2010), Dohmen et al. (2011), Dohmen and Falk (2011), Zhang (2012a). This is a promising and important approach to show the external validity of traits measured in the lab, but more importantly to show their economic significance, that is, confirm 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 administer the experiment while students still share the same experiences, several months before they make their education choice.

The paper most closely related to ours, and the only other one on 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. Using a structural approach, 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 surprisingly, 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 differences for ethnic minorities among high school children from the same area.

By validating the importance of competitiveness, our paper opens up new research questions. For example, how does competitiveness predict the performance of students in various study profiles? One could imagine that competitive students fare better in terms of grades than their less competitive peers. On the other hand, competitiveness may lead students to "overreach" and enter study profiles 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.

Future research will determine whether our result 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 result holds when this is not the case. Given our results, a perhaps more important question is what competitiveness exactly measures, and how it is correlated with other traits which may 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.[32]

References

Albrecht, James, Anders Björklund, and Susan Vroman, "Is There a Glass Ceiling in Sweden?," *Journal of Labor Economics*, 2003, *21*, 145–177.

Altonji, Joseph G. and Rebecca Blank, "Race and gender in the labor market," in Orley C. Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 3C, Amsterdam: Elsevier Science Publishers B.V., 1999, pp. 3144–3259.

[32]Furthermore, it is still far from clear to which extent gender differences in competitiveness are determined by nature and by nurture. Buser (2012) and Wozniak et al. (2010) find that for women the likelihood of entering the tournament varies over the menstrual cycle and Hoffman and Gneezy (2010) find that it is correlated with handedness. On the other hand, Gneezy et al. (2009) find that the gender gap in competitiveness varies between a patriarchal and a matrilineal society and Cardenas et al. (2012) find that gender differences in competitiveness vary across countries and may be correlated with gender stereotypes.

Arulampalam, Wiji, Alison L. Booth, and Mark L. Bryan, "Is there a glass ceiling over Europe? Exploring the gender pay gap across the wage distribution," *Industrial and Labor Relations Review*, 2007, *60*, 163–186.

Ashraf, Nava, Dean Karlan, and Wesley Yin, "Tying Odysseus to the Mast: Evidence From a Commitment Savings Product in the Philippines," *The Quarterly Journal of Economics*, 2006, *121* (2), 635–672.

Balafoutas, Loukas and Matthias Sutter, "Affirmative Action Policies Promote Women and Do Not Harm Efficiency in the Laboratory," *Science*, 2012, *335* (6068), 579–582.

Bertrand, Marianne, "New Perspectives on Gender," in Orley Ashenfelter and David Card, eds., *Handbooks in Economics Handbook of Labor Economics*, Vol. 4, part B of *Handbook of Labor Economics*, Elsevier, 2011, pp. 1543 – 1590.

_ **and Kevin F. Hallock**, "The Gender gap in top corporate jobs," *Industrial and Labor Relations Review*, 2001, *55* (1), 3–21.

_ **, 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*, 2011, *forthcoming.*

Bettinger, Eric P. and Bridget Terry Long, "Do Faculty Serve as Role Models? The Impact of Instructor Gender on Female Students," *The American Economic Review*, 2005, *95* (2), 152–157.

Borghans, Lex, Angela Lee Duckworth, James J. Heckman, and Bas ter Weel, "The Economics and Psychology of Personality Traits," *Journal of Human Resources*, 2008, *43* (4), 972–1059.

Brown, Charles and Mary Corcoran, "Sex-Based Differences in School Content and the Male-Female Wage Gap," *Journal of Labor Economics*, 1997, *15* (3), pp. 431–465.

Buser, Thomas, "The impact of the menstrual cycle and hormonal contraceptives on competitiveness," *Journal of Economic Behavior & Organization*, 2012, *83* (1), 1–10.

Cardenas, Juan Camilo, Anna Dreber, Emma von Essen, and Eva Ranehill, "Gender differences in competitiveness and risk taking: comparing children in Colombia and Sweden," *Journal of Economic Behavior & Organization*, 2012, *83* (1), 11–23.

Carrell, Scott E., Marianne E. Page, and James E. West, "Sex and Science: How Professor Gender Perpetuates the Gender Gap," *The Quarterly Journal of Economics*, 2010, *125* (3), 1101–1144.

Ceci, Stephen J., Wendy M. Williams, and Susan M. Barnett, "Women's underrepresentation in science: Sociocultural and biological considerations.," *Psychological Bulletin*, 2009, *135* (2), 218–261.

Croson, Rachel and Uri Gneezy, "Gender differences in preferences," *Journal of Economic Literature*, 2009, *47* (2), 448–74.

Cunha, Flavio and James Heckman, "The Technology of Skill Formation," *American Economic Review*, 2007, *97* (2), 31–47.

Cvencek, Dario, Andrew N. Meltzoff, and Anthony G. Greenwald, "Math-Gender Stereotypes in Elementary School Children," *Child Development*, 2011, *82* (3), 766–779.

Dee, Thomas S., "Teachers and the Gender Gaps in Student Achievement," *Journal of Human Resources*, 2007, *XLII* (3), 528–554.

Dohmen, Thomas and Armin Falk, "Performance Pay and Multidimensional Sorting: Productivity, Preferences, and Gender," *American Economic Review*, 2011, *101* (2), 556–590.

_ , _ , David Huffman, Uwe Sunde, Jurgen Schupp, and Gert G. Wagner, "Individual risk attitudes: Measurement, determinants, and behavioral consequences," *Journal of the European Economic Association*, 2011, *9* (3), 522–550.

Eccles, Jacquelynne, "Perceived Control and the Development of Academic motivation: Commentary," *Monographs of the Society for Research in Child Development*, 1998, *63* (2-3), 221–231.

Eckel, Catherine C. and Philip J. Grossman, "Sex Differences and Statistical Stereotyping in Attitudes toward Financial Risk.," *Evolution and Human Behavior*, 2002, *23* (4), 281–95.

Ellison, Glenn and Ashley Swanson, "The Gender Gap in Secondary School Mathematics at High Achievement Levels: Evidence from the American Mathematics Competitions," *The Journal of Economic Perspectives*, 2010, *24* (2), 109–128.

Felder, Richard, Gary Felder, Merdith Mauney, Charles Hamrin, and Jacquelin Dietz, *Journal of Engineering Education*, 1995, *84* (2), 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," Technical Report, NBER Working Paper No. w16546 2010.

Gneezy, Uri, Kenneth L. Leonard, and John A. List, "Gender Differences in Competition: Evidence From a Matrilineal and a Patriarchal Society," *Econometrica*, 2009, *77* (5), 1637–1664.

_ , Muriel Niederle, and Aldo Rustichini, "Performance In Competitive Environments: Gender Differences," *The Quarterly Journal of Economics*, 2003, *118* (3), 1049–1074.

Goldin, Claudia and Cecilia Rouse, "Orchestrating Impartiality: The Impact of 'Blind' Auditions on Female Musicians," *American Economic Review*, 2000, *90* (4), 715–741.

_ , **Lawrence F. Katz, and Ilyana Kuziemko**, "The Homecoming of American College Women: The Reversal of the College Gender Gap," *The Journal of Economic Perspectives*, 2006, *20* (4), 133–156.

Grogger, Jeff and Eric Eide, "Changes in College Skills and the Rise in the College Wage Premium," *The Journal of Human Resources*, 1995, *30* (2), 280–310.

Group, Inc Goodman Research, "Final Report of The Women's Experience in College Engineering Project," Technical Report, `http://www.grginc.com/WECE_FINAL_REPORT.pdf` 2002.

Hoffman, Moshe and Uri Gneezy, "Left Handed Women are More Competitive than Right Handed Men: On the Biological Basis of Gender Differences in Competitiveness," *unpublished manuscript*, 2010.

Huguet, Pascal and Isabelle Regner, "Stereotype Threat among School Girls in Quasi-Ordinary Classroom Circumstances," *Journal of Educational Psychology*, 2007, *99* (3), 545–560.

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," *The American Economic Review*, 2005, *95*, 1688–1699.

Kiefer, Amy K. and Denise Sekaquaptewa, "Implicit stereotypes, gender identification, and math-related outcomes: a prospective study of female college students.," *Psychological Science*, 2007, *18* (1), 13–18.

Kleinjans, Kristin J., "Do gender differences in preferences for competition matter for occupational expectations?," *Journal of Economic Psychology*, 2009, *30*, 701–710.

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*, 1992, *24* (3), 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*, 2010, *2* (1), 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*, 2000, *19* (4), 547–568.

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

_ **and** _ , "Gender and Competition," *Annual Review of Economics*, 2011, *3* (1), 601–630.

_ **, Carmit Segal, and Lise Vesterlund**, "How Costly is Diversity? Affirmative Action in Light of Gender Differences in Competitiveness," 2012.

Nosek, Brian A., Mahzarin R. Banaji, and Anthony G. Greenwald, "Math = male, me = female, therefore math ï¿œ me.," *Journal of Personality and Social Psychology*, 2002, *83* (1), 44–59.

Paglin, Morton and Anthony M. Rufolo, "Heterogeneous Human Capital, Occupational Choice, and Male-Female Earnings Differences," *Journal of Labor Economics*, 1990, *8* (1), 123–144.

Pautler, Emmanuelle, "The Links between Secondary and Higher Education in France," *European Journal of Education*, 1981, *16* (2), pp. 185–195.

Preckel, Franzis, Thomas Goetz, Reinhard Pekrun, and Michael Kleine, "Gender Differences in Gifted and Average-Ability Students," *Gifted Child Quarterly*, 2008, *52* (2), 146–159.

Roeder, Peter M. and Sabine Gruehn, "Geschlecht und Kurswahlverhalte," *Zeitschrift für Pädagogik*, 1997, *43* (6), 877–894.

Schroter Joensen, Juanna and Helena Skyt Nielsen, "Math and Gender: What if Girls do Math?," unpublished, Stockholm School of Economics 2011.

Segal, Carmit, "Working When No One is Watching: Motivation, Test Scores, and Economic Success," *Management Science*, 2012, *forthcoming.*

Sutter, Matthias and Daniela Rützler, "Gender Differences in Competition Emerge Early in Life," June 2010. IZA DP No. 5015.

Weinberger, Catherine J., "Mathematical College Majors and the Gender Gap in Wages," *Industrial Relations: A Journal of Economy and Society*, 1999, *38* (3), 407–413.

_ , "Is Teaching More Girls More Math the Key to Higher Wages?," in Mary C. King, ed., *Squaring Up: Policy Strategies to Raise Women's Incomes in the United States*, University of Michigan Press, 2001, chapter 11.

_ , "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, "Choices About Competition: Differences by gender and hormonal fluctuations, and the role of relative performance feedback," *unpublished manuscript*, 2010.

Zhang, Y. Jane, "Can Experimental Economics Explain Competitive Behavior Outside the Lab?," 2012.

_ , "The Communist Experiment in China: Narrowing the Gender Gap in Competitive Inclination," 2012.

Online appendix

Table 10. Means and standard deviations of control variables

	Mean	Standard-dev.
Math grade	6.63	1.07
GPA	6.89	0.62
Math relative	0.37	0.26
Math difficulty	3.80	2.80
Math quartile	2.12	0.96
Guessed rank	2.35	0.94
Risk	6.23	1.90
Lottery	3.22	1.32

Table 11. Gender and profile choice (alternative specifications)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	NT/NH as NT and ES/CS and ES			NT/NH and ES/CS as separate profiles			Students' own ranking		
Female	-0.275**	-0.318**	-0.210	-0.433***	-0.479***	-0.369***	-0.414***	-0.525***	-0.449***
	(0.121)	(0.127)	(0.130)	(0.115)	(0.120)	(0.122)	(0.117)	(0.126)	(0.128)
Math Grade		0.174	0.001		0.031	-0.151		-0.174	-0.322**
		(0.162)	(0.171)		(0.144)	(0.151)		(0.134)	(0.140)
GPA		0.248**	0.244**		0.193**	0.183**		0.234***	0.220**
		(0.102)	(0.101)		(0.095)	(0.093)		(0.088)	(0.087)
Rel. Math Gr.		-0.121	-0.072		-0.242**	-0.193		-0.341***	-0.316**
		(0.131)	(0.132)		(0.123)	(0.124)		(0.124)	(0.125)
Math Difficulty			-0.173*			-0.197**			-0.204**
			(0.089)			(0.085)			(0.089)
Math Quartile			-0.340***			-0.346***			-0.145*
			(0.079)			(0.075)			(0.075)
Cut 1	-1.475***	2.057	-0.015	-1.571***	0.319	-1.932	-1.505***	-0.640	-2.269*
Cut 2	-0.281***	3.341**	1.366	-1.299***	0.600	-1.636	-0.578***	0.349	-1.250
Cut 3	0.098	3.770***	1.835	-0.372***	1.603	-0.543	0.251***	1.241	-0.329
Cut 4				0.006	2.026*	-0.082			
Cut 5				0.568***	2.635**	0.576			
F/(Cmax-C1)	-0.175**	-0.186***	-0.113*	-0.202***	-0.207***	-0.147***	-0.236***	-0.279***	-0.231***
Observations	342	342	342	342	342	342	354	354	354

Dependent variable: Profile choice; coefficients are from ordered probit regressions; robust standard errors in parentheses; p-values for F/(Cmax-C1) are bootstrapped;
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 12. Profile choice: ordered probit regression (NT/NH as NT and ES/CS as ES)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Female	-0.279**	-0.210	-0.341***	-0.288**	-0.126	-0.075	-0.232*	-0.182	-0.234*	-0.193	-0.214	-0.164	-0.212	-0.174
	(0.123)	(0.128)	(0.131)	(0.135)	(0.128)	(0.134)	(0.134)	(0.139)	(0.134)	(0.138)	(0.136)	(0.139)	(0.136)	(0.139)
Entry		0.316**		0.284**		0.252*		0.280**		0.326**		0.335**		0.371**
		(0.137)		(0.137)		(0.140)		(0.142)		(0.155)		(0.154)		(0.166)
Math grade			0.177	0.151			0.005	-0.020	0.005	-0.017	0.017	-0.020	0.017	-0.017
			(0.163)	(0.164)			(0.172)	(0.172)	(0.173)	(0.172)	(0.173)	(0.172)	(0.173)	(0.171)
GPA			0.264***	0.286***			0.262***	0.283***	0.262***	0.284***	0.248**	0.269***	0.248**	0.270***
			(0.103)	(0.101)			(0.101)	(0.101)	(0.101)	(0.101)	(0.102)	(0.101)	(0.102)	(0.101)
Math relative			-0.113	-0.113			-0.070	-0.071	-0.070	-0.077	-0.048	-0.050	-0.048	-0.055
			(0.135)	(0.135)			(0.136)	(0.136)	(0.136)	(0.136)	(0.137)	(0.136)	(0.137)	(0.136)
Math difficulty					-0.272***	-0.264***	-0.160*	-0.156*	-0.160*	-0.156*	-0.175*	-0.176*	-0.175*	-0.176*
					(0.080)	(0.080)	(0.088)	(0.088)	(0.089)	(0.089)	(0.091)	(0.090)	(0.091)	(0.091)
Math quartile					-0.372***	-0.372***	-0.349***	-0.351***	-0.350***	-0.359***	-0.354***	-0.357***	-0.353***	-0.364***
					(0.083)	(0.084)	(0.082)	(0.082)	(0.082)	(0.084)	(0.081)	(0.082)	(0.082)	(0.083)
Guessed rank									0.007	0.072			-0.005	0.060
									(0.084)	(0.092)			(0.088)	(0.095)
Risk											-0.096	-0.140*	-0.096	-0.136*
											(0.073)	(0.076)	(0.073)	(0.076)
Lottery											0.114	0.104	0.114	0.103
											(0.075)	(0.075)	(0.075)	(0.076)
Cut 1	-1.249***	-1.202***	2.334*	2.458*	-2.547***	-2.497***	0.350	0.472	0.370	0.717	0.246	0.131	0.225	0.348
Cut 2	-0.041	0.021	3.636***	3.773***	-1.182***	-1.121***	1.749	1.885	1.770	2.131	1.651	1.551	1.631	1.769
Cut 3	0.346	0.411*	4.076***	4.215***	-0.723***	-0.662**	2.227	2.365*	2.248	2.612*	2.133	2.036	2.112	2.254
Female/(C3-C1)	-0.175**	-0.130*	-0.196***	-0.164**	-0.069	-0.041	-0.124**	-0.102*	-0.125**	-0.102*	-0.113*	-0.086	-0.113*	-0.091
Dif.	25.4%		16.4%		40.8%		22.5%		18.5%		23.4%		19.1%	
Bootstrap p-value	0.011		0.018		0.038		0.025		0.022		0.017		0.023	
Observations	342	342	342	342	342	342	342	342	342	342	342	342	342	342

Note: Coefficients are from ordered probit regressions, where NT>NH>ES>CS. All specifications include controls for performance in rounds 1 and 2 and the chance of winning. Coefficients are from ordered probit regressions; robust standard errors in parentheses; p-values for F/(C3-C1) and Dif. are bootstrapped; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The p-values for the impacts of confidence (comparing columns (7) and (9)) and risk attitudes (comparing columns (7) and (11)) on the gender gap (Female/(C3-C1)) are 0.53 and 0.24 respectively.

Table 13. Profile choice: ordered probit regression (NT/NH and ES/CS as separate profile)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Female	-0.441***	-0.373***	-0.502***	-0.450***	-0.305**	-0.254**	-0.392***	-0.341***	-0.400***	-0.359***	-0.358***	-0.312**	-0.364***	-0.327**
	(0.117)	(0.121)	(0.123)	(0.126)	(0.119)	(0.125)	(0.124)	(0.129)	(0.125)	(0.129)	(0.125)	(0.129)	(0.126)	(0.128)
Entry		0.317**		0.278**		0.252**		0.277**		0.338**		0.315**		0.366**
		(0.127)		(0.127)		(0.128)		(0.131)		(0.141)		(0.140)		(0.150)
Math grade			0.025	-0.006			-0.158	-0.187	-0.155	-0.185	-0.137	-0.175	-0.135	-0.172
			(0.145)	(0.145)			(0.152)	(0.152)	(0.152)	(0.151)	(0.151)	(0.150)	(0.151)	(0.149)
GPA			0.206**	0.227**			0.197**	0.218**	0.197**	0.220**	0.182*	0.201**	0.182*	0.203**
			(0.095)	(0.094)			(0.093)	(0.093)	(0.094)	(0.093)	(0.094)	(0.094)	(0.094)	(0.094)
Math relative			-0.240*	-0.243*			-0.198	-0.202	-0.200	-0.210*	-0.171	-0.174	-0.173	-0.182
			(0.126)	(0.126)			(0.127)	(0.127)	(0.126)	(0.126)	(0.128)	(0.127)	(0.128)	(0.127)
Math difficulty					-0.257***	-0.248***	-0.184**	-0.180**	-0.185**	-0.181**	-0.201**	-0.202**	-0.201**	-0.203**
					(0.075)	(0.075)	(0.084)	(0.083)	(0.084)	(0.084)	(0.086)	(0.085)	(0.087)	(0.086)
Math quartile					-0.372***	-0.373***	-0.354***	-0.357***	-0.357***	-0.367***	-0.360***	-0.363***	-0.362***	-0.372***
					(0.078)	(0.079)	(0.077)	(0.078)	(0.077)	(0.079)	(0.077)	(0.078)	(0.078)	(0.078)
Guessed rank									0.029	0.096			0.021	0.085
									(0.077)	(0.082)			(0.081)	(0.086)
Risk											-0.094	-0.135*	-0.091	-0.130*
											(0.067)	(0.069)	(0.067)	(0.069)
Lottery											0.155**	0.148**	0.155**	0.146**
											(0.072)	(0.071)	(0.072)	(0.071)
Cut 1	-1.332***	-1.290***	0.540	0.617	-2.631***	-2.587***	-1.647	-1.569	-1.552	-1.233	-1.624	-1.762	-1.541	-1.446
Cut 2	-1.059***	-1.012***	0.821	0.903	-2.340***	-2.291***	-1.349	-1.266	-1.254	-0.930	-1.321	-1.454	-1.238	-1.139
Cut 3	-0.118	-0.059	1.840*	1.933*	-1.263***	-1.204***	-0.242	-0.147	-0.148	0.190	-0.198	-0.317	-0.115	-0.001
Cut 4	0.268	0.330	2.273**	2.366**	-0.807***	-0.747***	0.228	0.324	0.323	0.663	0.275	0.159	0.358	0.476
Cut 5	0.841***	0.906***	2.893***	2.988***	-0.147	-0.086	0.897	0.995	0.993	1.336	0.948	0.835	1.032	1.153
Female/(C5-C1)	-0.203***	-0.170***	-0.213***	-0.190***	-0.123***	-0.102**	-0.154***	-0.133***	-0.157***	-0.140***	-0.139***	-0.120***	-0.141***	-0.126***
Dif.	16.2%		11.0%		17.2%		13.5%		11.2%		13.7%		11.0%	
Bootstrap p-value	0.007		0.015		0.026		0.018		0.011		0.014		0.016	
Observations	342	342	342	342	342	342	342	342	342	342	342	342	342	342

Note: Coefficients are from ordered probit regressions, where NT>NT/NH>NH>ES>ES/CS>CS. All specifications include controls for performance in rounds 1 and 2 and the chance of winning. Coefficients are from ordered probit regressions; robust standard errors in parentheses; p-values for F/(C3-C1) and Dif. are bootstrapped; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The p-values for the impacts of confidence (comparing columns (7) and (9)) and risk attitudes (comparing columns (7) and (11)) on the gender gap (Female/(C3-C1)) are 0.64 and 0.08 respectively.

Table 14. Profile choice: ordered probit regression (students' own ranking)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Female	-0.384***	-0.316**	-0.503***	-0.447***	-0.299**	-0.243*	-0.423***	-0.366***	-0.420***	-0.377***	-0.399***	-0.357***	-0.402***	-0.369***
	(0.118)	(0.124)	(0.128)	(0.133)	(0.120)	(0.126)	(0.130)	(0.134)	(0.132)	(0.135)	(0.132)	(0.135)	(0.134)	(0.136)
Entry		0.300**		0.280**		0.252*		0.288**		0.324**		0.253*		0.288*
		(0.136)		(0.139)		(0.136)		(0.139)		(0.148)		(0.147)		(0.155)
Math grade			-0.220	-0.249*			-0.374***	-0.404***	-0.374***	-0.404***	-0.369***	-0.396***	-0.369***	-0.397***
			(0.137)	(0.137)			(0.142)	(0.142)	(0.142)	(0.142)	(0.142)	(0.142)	(0.142)	(0.142)
GPA			0.225**	0.247***			0.213**	0.235***	0.213**	0.236***	0.220**	0.234***	0.220**	0.237***
			(0.088)	(0.087)			(0.087)	(0.087)	(0.087)	(0.087)	(0.088)	(0.088)	(0.088)	(0.088)
Math relative			-0.372***	-0.372***			-0.351***	-0.351***	-0.350***	-0.356***	-0.355***	-0.351***	-0.356***	-0.357***
			(0.126)	(0.126)			(0.127)	(0.128)	(0.127)	(0.127)	(0.127)	(0.127)	(0.127)	(0.127)
Math difficulty					-0.233***	-0.222***	-0.186**	-0.180**	-0.186**	-0.181**	-0.178**	-0.178**	-0.178**	-0.179**
					(0.077)	(0.077)	(0.091)	(0.089)	(0.091)	(0.090)	(0.089)	(0.089)	(0.089)	(0.089)
Math quartile					-0.188**	-0.193***	-0.172**	-0.179**	-0.171**	-0.185**	-0.174**	-0.179**	-0.175**	-0.186**
					(0.074)	(0.074)	(0.075)	(0.075)	(0.076)	(0.076)	(0.075)	(0.075)	(0.075)	(0.076)
Guessed rank									-0.009	0.055			0.010	0.059
									(0.081)	(0.087)			(0.082)	(0.087)
Risk											0.057	0.024	0.059	0.029
											(0.067)	(0.068)	(0.067)	(0.068)
Lottery											0.048	0.039	0.048	0.038
											(0.065)	(0.066)	(0.065)	(0.066)
Cut 1	-0.841***	-0.796***	-0.472	-0.370	-1.675***	-1.627***	-2.122*	-2.024	-2.148*	-1.841	-1.733	-1.821	-1.693	-1.609
Cut 2	0.105	0.156	0.530	0.639	-0.671***	-0.618**	-1.087	-0.982	-1.113	-0.797	-0.694	-0.778	-0.654	-0.564
Cut 3	0.951***	1.009***	1.434	1.550	0.238	0.296	-0.154	-0.041	-0.180	0.144	0.242	0.163	0.282	0.377
Female/(C3-C1)	-0.215***	-0.175***	-0.264***	-0.233***	-0.157***	-0.127**	-0.215***	-0.185***	-0.214***	-0.190***	-0.202***	-0.180***	-0.204***	-0.186***
Dif.	18.4%		11.6%		19.1%		14.1%		11.1%		11.0%		8.9%	
Bootstrap p-value	0.014		0.022		0.031		0.020		0.016		0.045		0.037	
Observations	354	354	354	354	354	354	354	354	354	354	354	354	354	354

Note: Coefficients are from ordered probit regressions. All specifications include controls for performance in rounds 1 and 2 and the chance of winning. Coefficients are from ordered probit regressions; robust standard errors in parentheses; p-values for F/(C3-C1) and Dif. are bootstrapped; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The p-values for the impacts of confidence (comparing columns (7) and (9)) and risk attitudes (comparing columns (7) and (11)) on the gender gap (Female/(C3-C1)) are 0.46 and 0.13 respectively.