Persistency in Teachers' Grading Biases and Effect on Longer Term Outcomes: University Admission Exams and Choice of Field of Study [*]

March 19, 2017
Preliminary and Incomplete

Victor Lavy
University of Warwick, Hebrew University and NBER

Rigissa Megalokonomou
The University of Queensland

Abstract

Recent research focus on what shapes gender differences in academic achievements and in university field of study. In this paper we focus on how teachers' gender role attitudes and stereotypes influence the gender gap by affecting the environment at school. We explore the extent to which teachers' gender bias in high school influences students' academic performance in high-stake exams that determine admission to universities and on students' choice of university field of study. We use data from large number of high schools in Greece and measure teachers' bias as the difference between a student's school exam score in 11th and in 12th grade (scored 'non-blindly' by the students' teachers) and her national exam score (taken at the end of 11th and 12th grade and scored blindly). We then define a teachers' bias measure at the class level by the difference between boy's and girl's average gap between the school score and the national score. Positive values indicate that a teacher is biased in favor of boys in a particular subject. We link teachers over time and are therefore able to get a persistent teacher's bias measure based on multiple classes, and the effect is estimated for later students' performance. The panel data on teachers relieves concerns that our measure of gender bias may just pick up random (small sample) variation in the unobserved "quality" or "non-cognitive" skills of the boys vs. girls in a particular single class or any other class specific dynamics. Our results may be summarized with three broad conclusions. First, the same teachers who are biased for one class are biased in the same way for other classes in the same year and in classes in earlier or later academic years. The very high correlations of within teachers' biases in different classes reveal high persistency in teachers' stereotypical behavior. Second, teachers' biases in core and elective subjects (classics, social science, science, exact science) have positive effect on boys' and negative effects on girls' performance on end of high school university admission exams. Female teachers are more pro-girls on average but the effect of female and male teachers' biases on national exams are not statistically different. Third, teachers' biases in specific courses lower the likelihood that students enroll in a related field of study at the university. This average effect masks large heterogeneity by gender, being larger and statistically significant for girls and not different from zero for boys. However, the effect on choice of STEM subjects are large and positive for boys and small and insignificant for girls.

[*] We thank participants at the brown bag seminars at the Hebrew University, The University of Queensland, and University of Warwick for useful discussions and suggestions. The second author acknowledges financial support from the European Research Council through ERC Advance Grant 323439 and from CAGE at the Department of Economics at the University of Warwick.

1. Introduction

A robust stylized fact established in recent years in many countries show that girls out-perform boys in school achievements in primary and secondary school. The gap is larger in school tests that are graded by the school teachers and smaller in external exams that are graded 'blindly'. The gaps are smaller in STEM subjects, often showing still boys' advantage followed universally by higher college enrollment rates of men in these fields of study. For example, the National Center for Educational Statistics (NCES) 2015 report show that 57 percent of all Bachelor's degrees conferred by postsecondary institutions in the US in 2013-14 were to women while in STEM subjects it was much lower, 39 percent in physical sciences and science technologies, 18 percent in computer and information sciences, 18 percent in engineering and engineering technologies, and 10 percent in computer engineering.[1] This skewed pattern of gender differences in college fields of study naturally determines gender occupational differences in the labor market. For example, only 14% of engineers in the US are women, though this rate is much higher than in the early 1980s, when only 5.8% of engineers in the U.S. were women.[2]

The debate of what shapes these gender differences in academic achievements and in university fields of study is the focus of much recent research.[3] In this paper we focus on how teachers' gender role attitudes and stereotypes influence the gender gap by affecting the environment at school. We explore the extent to which teachers' gender bias in high school influences students' academic performance in high-stake exams that determine admission to universities and on students' choice of university field of study. We use data from large number of high school in Greece where the performance in these high stake exams are the sole determinant of university admission. Our sample includes female and male teachers, so the analysis reflects potential bias due to the teacher's own sex which allow us to distinguish the gender bias by teachers gender. We measure teachers' bias as the difference between a student's school exam score in 11th and in 12th grade (scored by the student's teacher) and his or her external exam score (taken at the end of 11th and 12th grade and scored nationally).[4] We then define a teachers' bias measure at the

[1] The female share in degrees conferred in health professions and related programs was 84 percent, in English language and literature/letters was 69 percent, in biological and biomedical sciences was 58 percent, in mathematics and statistics 43 percent. https://nces.ed.gov/programs/digest/d15/tables/dt15_318.30.asp?current=yes.

[2] STEM Education: Preparing for the Jobs of the Future, A Report by the Joint Economic Committee Chairman's Staff Senator Bob Casey, Chairman April 2012.

[3] Some studies emphasize the role of biological gender differences in determining gender cognitive differences (Witelson 1976, Lansdell 1962, Waber 1976), while others emphasize the social, psychological and environmental factors that might influence this gap. There is limited credible evidence for this debate because it is difficult to disentangle the impact of biological gender dissimilarities from environmental conditions, and because it is difficult to measure stereotypes and prejudices and test their causal implications.

[4] The systematic difference between non-blind and blind assessment across groups as a measure of discrimination or stereotypes was pioneered in economics by Blank (1991) and Goldin and Rouse (2000). This approach was first applied to the economics of education in Lavy (2008), to measure gender bias in grading by teachers and it was followed by others, for example, Björn, Höglin, and Johannesson (2011), Hanna and Linden (2012), Cornwell,

class level by the difference between boy's and girl's average gap between the school score and the national score. Positive values indicate that a teacher is biased in favor of boys in a particular subject. We link teachers over time and are therefore able to get a persistent teacher bias measure based on multiple classes (on average 6 classes per teacher), and the effect is estimated for later student performance. The panel data on teachers relieves concerns that our measure of gender bias may just pick up random (small sample) variation in the unobserved "quality" or "non-cognitive" skills of the boys vs. girls in a particular single class or any other class specific dynamics. We find that the same teachers who are biased for one class are biased in the same way for other classes in the same year and in earlier or later academic years. This evidence of a "persistent" (average) teacher bias component (across multiple years/classes) is reassuring that our bias measure is not picking up random variation in the mix of boys and girls on unobserved attributes.

We use data of high school teachers and students in Greece for the period 2003-2011. Our panel data of teachers includes 900 teachers from 21 high schools over this entire period. Using this sample, we find that the bias measures derived from the teachers' panel data yield results very similar to those obtained when measuring the bias based only on current own class. Perhaps this result is expected given the higher correlation and persistence of teacher's biases measured in different classes. We also have a sample of an additional 116 schools for which we do not have panel data on teachers. Using this sample we measure teachers' biases based on grading information of their current own classes and obtain estimates that are similar to the respective results obtained from the 21 schools sample.

We first estimate the effect of the bias of teachers in 11^{th} grade on students' performance in the national exams at the end of 12^{th} grade. We measure the bias in each subject and then average them over bundles of subjects as follows. The first bundle includes core subjects that all students have to study, including Modern Greek, history, physics, algebra and geometry. The other three bundles are the three study tracks available to students in 11^{th} and 12th grade, classics, science and exact science. In 11^{th} grade the subjects taught in the classics track include ancient Greek, philosophy and Latin, in the science track they are mathematics, physics, and chemistry, and in the exact science track they include mathematics, physics and computer science. In 12^{th} grade the subjects taught in the classics track are ancient Greek, Latin, literature and history, in the science track they are biology, mathematics, physics and chemistry, and in the exact science track they are mathematics, physics, business administration and computer science. We find that the teachers' biases in all four groups of subjects have positive effect on boys and

Mustard, Van Parys (2013), Burgess and Greaves (2013), and Botelho, Madeira and Rangel (2015), who implemented the same methodology using data from other countries and getting overall similar evidence about teachers' stereotypes/biases.

negative effects on girls 2th grade external exams scores. All the estimates are precise and statistically significantly different from zero except in the classics track. Based on bias measures derived from the teachers' panel data, the effects size (in terms of standard deviation of the test score distribution) in the core subjects is 0.103 for boys and -0.111 for girls, in classics they are 0.067 and -0.067 respectively, in science they are 0.128 and -0.088 and in exact science 0.051and -0.154. We find that the effect of female and male teachers' biases are not statistically different except for the effect in the exact science track.

The psychology and sociology literature provide ample evidence about the potential mechanisms of the effect of teacher's gender stereotypical attitudes on students cognitive and non-cognitive outcomes. For example, teachers are said to treat the successes and failures of boys and girls differently, by encouraging boys to try harder and allowing girls to give up (Dweck et al. (1978) and Rebhorn and Miles (1999)). Sadker and Sadker (1985) suggest that teachers give more attention to boys by addressing them more often in class, giving them more time to respond and providing them with more substantive feedback. Teachers are also found to treat boys and girls differently, in particular with regard to math instruction: Hyde and Jaffe (1998) show that math teachers tend to encourage boys to exert independence by not using algorithms and that boys who pursue this rebellious approach are seen as having a promising future in mathematics; girls, on the other hand, are controlled more than boys, and are taught mathematics as a set of rules or computational methods. Leinhardt, Seewald and Engel (1979) find that teachers spent more time training girls in reading and less time in math, relative to boys. In addition, according to the National Center of Education Statistics (1997) girls are less likely than boys to be advised, counseled and encouraged to take courses in math.

In the second part of the paper we estimate the effect of 11th and 12th grade teachers' biases on university enrollment by field of study. Girls have higher enrollment rate in humanities departments, 33.4 percent versus 11 percent of boys. In social science and science, enrollment is not different by gender, 28 percent among boys and also among girls. In the exact sciences there are large enrollment disparities by gender, 11 percent among girls and 30 percent among boys. Based on multiple choice regressions with student's fixed effect, we find that 11th and 12th grade teachers' biases have a negative and statistically significant effect on girls' choice of program of study. A one standard deviation increase in teacher's bias in 11th grade lowers the probability of choosing the same field of study in university schooling by 13 percentage point. The effect of teacher's bias in 12th grade lowers this probability by 5 percentage point. The respective estimated effects on boys are positive but small and imprecise, practically not being different from zero.

This paper makes a substantive contribution to the literature on gender differences in STEM majors and careers by linking quantitative measures of teacher bias measures of children's later academic

outcomes. To our knowledge, this is the first paper to establish a believable causal connection between high school "culture" and the prevalence of gendered outcomes. Two earlier papers examined the effect of teachers' bias in primary schools on students' cognitive performance. Lavy and Sand (2015) analyses teachers' bias in primary schools in Tel Aviv, Israel, and estimate its effect on boys' and girls' test scores in math, English and Hebrew in middle school and high school tests and on choices about the level of study of math and science courses that they select in high school. This earlier study measures teachers' bias as we do in this paper, assessing teachers' biases towards one of the sexes, as reflected by a more positive evaluation on "non-blind" tests relative to the "blind" tests of this group. Their findings are similar to those reported in this paper, suggesting that teachers' biases favoring girls have a positive effect on girls' achievements and negative effect on boys' and vice versa and also impact students' enrollment in advanced level math courses in high school – girls positively and boys negatively. However, in this paper we are able based on panel data on teachers in high schools in Greece to assess the impact of the persistence component in teachers' stereotypical biases. Terrier (2015) estimate the effect of teachers' bias similarly to Lavy and Sand (2015), and her study also lack panel data on teachers that is needed to measure teachers' biases out of sample. She finds that in primary schools in France there is a positive correlation between teachers' grading bias in favor of boys in a specific subject and the progress of boys relative to girls in class in that subject.

The rest of the paper is organized as follows. In Section 2, we present our data. Section 3 explains the identification and estimation methodologies. We detail our results in Section 4, and Section 5 offers conclusions and policy implications.

2. Context and Data

2.1 The Greek Universities Admission System

University admission in Greece is based on a centralized process, administered by the Hellenic Ministry of Education. Universities in Greece are public, free with no tuition fees and admission is based on the national high school exit exams. Most undergraduate degrees in Greek universities take 4 years to complete on time, except of Polytechnic University in Athens (the most prestigious university among engineering departments) which takes 5 years to complete on time. Students applying to universities have to participate in standardized national tests for university admission. All schools that administer these tests follow the same curriculum and offer courses in core and track subjects in accordance with the material covered in the national exams. From 2006 onwards students take national exams only at the end

of 12th grade[5] and university admission is based on students' performance in these national exams and school exams that students take throughout the year. Until 2005 the average score used for university admission was also based on national exams taken at the end of 11th grade.[6]

The data we use in this study include the school and the national exams' test scores for all students. The national exams scripts are centrally collected by the Ministry of Education and are sent to examiners across the country while the name of the student and its gender are disguised. Therefore we denote the national exam scores "blind" scores as the external examiner does not know the name or the gender of the student. The school scores are based on school exams graded by the student's teacher and therefore they are 'non-blind, the identity and gender of the student are obviously concealed in this setting. Each student receives a report card at the end of each term that lists her/his test score in the school exam in each subject.[7] We use these non-blind test scores in each subject to examine teacher biases. Most students have different teachers in each subject while some might have the same teacher in two or more related subjects in the same year or grade. For the period 2003-2005 we observe the blind and non-blind scores in both grades (11th and 12th). From 2006 to 2011 the relevant data is the 12th grade national and school test scores.

Even though every student has a first and second term school test score in every subject, we prefer to use the latter because the second term exam is likely to cover the same material included in the end-of-year national exam and because the second term school exam and the national exam in the same subject are administered around the same time.[8] Schools often administer the school exam very much towards the end of the school year and before the national exam period. From the student's perspective both exams are high-stakes: the final grade in a given subject for university admission purposes is a weighted average of the national exam (70%) and the school exam (30%). The blind and non-blind scores are also important for grade completion, for obtaining a high school graduation diploma and also for consideration of drop-out and grade repetition decisions. Both scores are reported in the high school graduation diploma which is some time is by requested by employers.

[5] Which is the senior year of high school in Greece.

[6] The 11th grade national exam were given a small weight if the test's score was higher than the same subject exam of 12th grade.

[7] The non-blind scores could potentially be affected by a student's performance in previous class exams in the same term, if there are more than one class exams.

[8] However, we note here that we obtain very similar results when we use the first term non-blind scores. These results are not reported in this paper and are available from the authors.

The allocation of students to classes is by alphabetical order of surname.[9] The average number of classes in 11th and 12th grades is 4. Average number of students per class is 21 in 11th grade and 19 in 12th grade (Table 2). Therefore assignment of students to class and class peers composition are random and as a result classroom composition by ability and other peer characteristics should not vary within school and grade. It also means that students cannot choose a teacher based on observed or unobserved characteristics.

University applicants submit to the Ministry of Education a list of their preferred universities in a rank order. The application must include also the preferred field (department) of study.[10] The average score cutoffs for admission at different departments and universities is not known to students when submitting their application. However, students can apply to several university departments conditional on their high school track of study. At the beginning of the 11th and 12th grades, students enroll in a specific study track. There are three main tracks: classics, science and exact science.[11] All schools offer these three tracks and students choose one of them in 11th grade are unlikely change it until graduating from high school. Each track includes different subjects and all students in a track take the same school exams. It is conceivable that students' choice of track takes into consideration their aspiration for university field of study because of pre-requisite in admission to various university programs. For example, admission to an engineering school requires high school graduation in the science or exact science track. In addition to the track subjects, students have also to take exams in core compulsory subjects.

2.2 Data

In this study we use combined information obtained from schools and from administrative sources for a large number of high schools in Greece. The baseline sample is 11th grade students in 2003-2005 in 135 schools and 12th grade students in 2003-2011 in same schools. The sample includes 1,244 11th grade classes and 3,787 12th grade classes. The data we obtained from these schools' administrative records include information about students and teachers. The student level information includes identifiers for students and their classes (including class size in both grades), gender, year of birth, track of studies in high school, absenteeism records in 11th and 12th grade, drop out and repetition status in 11th and 12th grade, test scores of the school and national exams in all subjects in 11th and 12th grade. The

[9] There were only few schools during the interviews that reported that they assign students to classes based on the second language that they choose. The first compulsory language is English and the students chose mainly French or German.

[10] See Goulas and Megalokonomou (2016) for more details about the admission algorithm.

[11] This track is called in Greek "Information Technology".

sample includes public, private and experimental[12] schools, in large and smaller cities, in urban and rural areas (see map in Figure 1). We also obtain teachers and principals level information from administrative records of 21 of the high schools in our sample. The teachers' information permits tracking teachers through their teaching history during 2003-2011. We therefore constructed a panel data on teachers and school principals and matched them to their classes and students by year and subject for the whole period. We also conducted a short interview with the principals of these 21 schools. The information we obtained include the gender of teachers and principals.

We then link the students and teachers data sets with administrative data that we obtained from the Ministry of Education. The Ministry of Education collects regularly data on all students and all schools that participate in the national exams on which university admission is based. The ministry data cover all 12th grade students and it includes exactly the same information that we collected directly from the 135 schools. This dataset includes for each student the test scores in mathematics, physics, history, Modern Greek and physics in 12^{th} grade as well as the test scores in all subjects that are part of the student's high school study track. The raw exam score is at a 1-20 scale and we transform it into z-scores for each year, by type of exam and subject, to facilitate comparison over time and interpretation of our findings. In addition to this data, we also obtained from the Ministry of Education information about students' university enrollment and program of studies. This administrative data includes a student level university admission score computed by the Ministry based on the national and school exams scores. In addition this data provide information on the number of university applications submitted by each student, the name of the university where he enrolled, her/his field and degree of study.

Using data for the universe of students at each institution for every year, we compute for each higher education institution and for each study program (department) the admission test score cutoff which we use as the threshold for admission.

The Ministry data can be matched with the schools' data based on detailed information about students such as year of birth, gender, the high school attended, graduation year, track of high school study program, and the test scores in the national exams. We are able to match all students uniquely because of the very detailed level of each of these variables. Table 1 presents descriptive statistics for the sample of 135 schools. The proportion of female students is around 56 percent in both grades. The average GPA in 11^{th} and 12^{th} grade is 72 and 77, respectively. 92 percent of students attend public schools, 4 percent and 4 percent attend private and experimental schools, respectively. 90 percent are in

[12] These are public schools. Admission to these schools is based on a lottery for the years that we use in this study. In 2013 the admission process changed and students gain admission based on their performance in very competitive admission exams

urban areas. Almost 82 percent of students enroll eventually in a university. Students apply on average to 25 different combination of universities and departments[13] and on average they study their 8th most preferred university department. The proportion enrolled in exact science, science, humanities and social science departments in 2003-2011 are 15 percent, 4 percent, 19 percent and 22 percent, respectively.

Table 2 presents mean differences between the sample of 116 schools and the sample of 21 schools. The average number of classes in the first sample is 3.90 and in the second 3.92. Average class size is 18-20 students in 11th grade and 19-20 students in 12th grade. 37 percent of the students study in the classics track, 28 percent in the science track, and 43 in exact science track. The differences between the two samples are small and for some variables they are statistically insignificant.

3. Methodology and Estimation Framework

3.1 Measuring Teachers' Biases

The national exam scripts are centrally collected by the Ministry of Education and are assigned for grading to teachers in other schools while the name of the student and gender are concealed. The student id that appears in the first page does not reveal any information about the student. Therefore we denote the national exam scores as "blind". The school exams are graded by the class teacher and therefore the name, gender and other student's information are known to the examiner. We therefore denote the school score as "non-blind". For each student we observe a set of both national (blind) and school (non-blind) exams scores. Most students have different teachers in each subject while some might have the same teacher in two or more related subjects in the same year and grade. For the years 2003-2005 we have the data for the blind and non-blind scores for each of the 11th and 12th grade students. Starting in 2006, national exams were administered only 12th grade.

Table 4 presents the means of the blind and the non-blind scores, for boys and girls, and the gender differences between these test scores in 11th grade for 2003-2005. The gender gap varies by subject and type of exam. Boys outperform girls in the following subjects in the blind exams: physics, geometry and algebra (core subjects). In all other subjects girls outperform boys in the blind exams. The gender difference in the non-blind exams scores is always in favor of girls. In other words, girls always outperform boys in the non-blind exams which are graded by the class teacher. This girls' advantage is evident even in subjects where boys outscore girls in the external blind exams. These systematic gender

[13] Which is equivalent to submitting 25 degree applications.

differences are interesting as they imply that the achievement gap is always in favor of girls in the non-blind exams that are graded by the classroom teacher.

Table 5 presents the same descriptive tests statistics for the 12th grade students. The gender gaps in 12th grade have the same pattern as in 11th grade. Boys outperform girls in in the blind exams in mathematics and physics (core subjects) and in mathematics, physics and chemistry (science track). The differences between boys' and girls' blind scores are statistically significant in most cases and they vary from 0.40 in favor of girls in Modern Greek to 0.16 in favor of boys in physics (science track). However, girls obtain higher scores in the non-blind exam in all subjects. The non-blind - blind gender differences vary from 0.48 in favor of girls in Modern Greek to 0.055 in science track biology in favor of girls. The positive achievement gaps that boys face in the blind exams in mathematics and physics are not present in the non-blind exams graded by the classroom teachers.

We construct the teacher bias measure in two steps based on each student test scores in the blind and non-blind score by subject. We first compute for each student in each exam the difference between her/his non-blind and blind exam scores. We then average these differences for boys and for girls in each class and then compute the difference of this two means in each class. Stating differently, we define a teacher bias measure at the class level by the difference between boys' and girls' average gap between the non-blind score (NB) and the blind score (B):

$$\text{Teacher } j \text{ Gender Bias in Class } c = Mean_c\,[\Sigma_{ic}\,(NB_i - B_i\,|Male_i)] - Mean_c\,[\Sigma_{ic}\,(NB_i - B_i\,|Female_i)]$$

We repeat this procedure for every class, subject and grade. This measure takes negative and positive values depending on teacher' stereotypical behavior. Positive (negative) values indicate that a teacher is biased in favor of boys (girls) in this particular subject.

Since we have panel data for teachers' by class, subject, and year, we can compute the persistent part of a teacher stereotypical bias by averaging the bias measure over all of the teacher's classes during the study period. We do however want to exclude from this average the bias in the class in whom we want to estimate the impact of the teacher bias. Therefore, we construct the average bias of a teacher based on all her/his other classes except his current class. In other words, we measure the bias relevant for a particular classroom using outcomes from all other classes taught by the same teacher. For example, we measure a teacher's bias in classes in earlier years, other classes in the same year, other classes in later years and based all of these classes. The later measure uses all possible information about the teacher and therefore it is the one that reflects more reliably her/his persistent gender biased behavior. Following this approach of using 'out of sample' data to measure teachers' biases, we alleviate the concern that our teacher bias measure picks up class level unobserved variation in boys' and girls' behavior or other gender

differential non-cognitive characteristics. In Figure 3 we present the distribution of the teachers' bias measures, the first based on current own class test scores and the second based on the teacher's all other classes during the study period.

We present in Table 6 descriptive statistics about the number of classes taught on average by teachers in our 11th and 12th grade schools sample. We drop from the sample teachers that we observe only in one class in 11th or in 12th grade, because this mean that we do not observe them teaching other classes. An 11th grade teacher appears in the data over all years with more than 7 different classes. The average per year is 3.4. The average number per teacher of different modules is 1.6 and the average number of classes per year is 1.75. There is little variation in these statistics during 2003-2005 and on average teachers are present in our sample in two of these three years. A 12th grade teachers appears in our sample 10.8 times over the period 2003-2011, on average in 4.4 years. Otherwise the other statistics are similar to those of the 11th grade teachers.

In Table 7, columns 2-3, we present using the sample of the 21 schools the mean and standard deviation of the measure of teacher bias based on all other classes, by subject and grade. On average teacher are biased in favor of girls across all subjects. The bias in 11th grade is highest in computer science (-0.223) and lowest in physics (-0.004). Among 12th grade teachers, the bias is highest in physics (-0.231) and lowest in in ancient Greek (-0.060). In columns 4-5 we report the value of teachers' biases when they are measured based only the class they teach in a particular year. Here again teachers are on average pro-girls in all subjects, and the measures in different subjects are very similar to those based on using all other classes. The correlation coefficient between the two measures of bias are high, ranging from 0.60 to 0.86. These correlation coefficients for all subjects are presented in column 8 of Table 7. Their high values are evidence of the high persistency in teachers' gender bias behavior. In Table 8 we present additional evidence about the high correlation between the bias measure based on all other classes and the bias measure based on own current class. In the first panel we present the results based on the full sample of teachers, in the second panel we present the results for female teachers and in the third panel for male teachers. The pattern is clear, high persistency in teacher gender grading bias among female and among male teachers, the degree of persistency being very similar for the two genders.

Using the sample of 116 we are able to measure the teacher biases based on current own classes and compare them to the respective estimates obtained from the sample of 21 schools. These results are presented in online appendix Table A1 and show that the differences between the two sets of bias estimates are small and follow the same pattern across subjects. The combined evidence presented in Tables 6 and A1 imply that the estimates of teachers bias derived from the sample of 21 schools based on all other classes are very similar to the bias measures obtained from the sample of 116 schools based on

the current teachers' class. In the next section we report the results of estimating the effect of various measures of teachers' biases on students' short term academic performance (subsequent national exam tests, and other education outcomes-drop out and grade repetition), and on longer term outcomes, in particular on the choice of university field's study and on the quality of the institution.

4. Effect of Teacher Biases on High School Outcomes

We estimate the following model for obtaining the effect of teachers' biases in 11th grade on the performance of students in 12th grade national exams based on which universities determine admission:

$$Y_{icjt} = \alpha + \mu_c + \Theta_j + \lambda_t + \gamma X_{icjt} + \pi TB_{cj} + \varphi_{cj} + \psi_{icjt} \qquad (1)$$

where Y_{icjt} denotes the outcome of student *i*, in high school or class *c*, subject *j* and year t; X_{icjt} are the student characteristics which includes high school track of studies and the score in the national exam in subject *j*; μ_c is a high school or class fixed effect; Θ_j is a subject fixed effect; λ_t is a year fixed effect; TB_{cj} is the measure of teachers' biased behavior in school (class) c and subject *j*. The error term in the equation includes a school (or class) and subject specific random element φ_{cj} that allows for any type of correlation within observations of the same school across classes and an individual random element ψ_{icjt}. The coefficients of interest is π and it captures the effect of teacher's biases on academic outcomes.

In Table 9 we present the results of estimating equation (1) to obtain an effect of 11th grade teacher bias measured in all other classes on the blind score in 12th grade using the sample of 21 schools. We present estimates from three different regression specification: the first includes subject and year fixed effects, in the second we add school fixed effects and the third includes a class instead of a school fixed effect. Standard errors are clustered at the school level. The estimated effects in all three specifications are positive in the boys' regressions and negative in the girls' regressions. The boys' regressions are somewhat sensitive to adding the various controls, while in the girls' regressions the estimates are almost identical in all three specifications. In the core, classics and science subjects regressions the boys' and the girls' estimated coefficients are almost identical but with an opposite sign. A one sd increase in 11th grade core subjects teacher bias will increase the boys test score in 12th grade by 0.10 sd and will reduce the girls test score by 0.11 sd. In science the respective effects are larger, a 0.13 increase among boys and a 0.09 decrease among girls. In exact science the effect is large negative for girls, 0.154 while for boys it is small and not different from zero.

Table A6 in online appendix is a mirror image of Table 8 but the bias measure is derived from the teacher current year class. Remarkably, the point estimates in panel A of this table obtained from the

sample of 22 schools depict the similar pattern and similar point estimates. It is striking that the bias estimate measure derived from a sample that includes on average only 25 students in the teacher's current year class yields very similar point estimates to those obtained from a bias measure obtained as an average of the teacher's bias in 7 other classes taught in same, earlier or later years. In panel B of the same table we estimate the effect of same type of bias measure using the sample of 135 schools. These estimates are very similar to those presented in panel A and to those presented in Table 9.

Summarizing the evidence presented in Tables 7-9, we find that persistent teachers pro-boys bias in 11th grade has positive effect on boys and negative effect on girls' test score in the 12th grade national exam. The absolute effect size is similar by gender and they are economically meaningful. We find that measuring the bias based on current class with less than 30 students yields very similar estimates to those obtained when the bias is measured based on over 7 class with over 200 other students in the current or other years. These results suggest that the various bias measures are highly correlated. In Table 8 we present the correlation coefficients between the persistent teacher bias measure, either based on other classes in same year or in other class in any year, and the measure based on current class. These correlation coefficient estimates are positive and large in the sample of all teachers (panel A) and in the female teachers' sample (panel B) and the male teachers' sample (panel C). For example, the correlation between the 11th grade bias measure in other classes in any year and the bias measure in current own class in 11th grade is 0.813 when estimated in a regression with only subject and year fixed effect. When using the bias measured in other classes in any year the estimate is very similar, 0.720. The high degree of persistence is similar to male and female teachers.

Treatment Effect Heterogeneity by Female and Male Teachers and School Principals

In Table 10 we present estimates when we allow the effect of teacher's bias to vary the teacher's gender. We add to equation (1) and interaction term between TB_{cj} and an indicator for female teacher and we also include in the equation a main effect for teacher's gender. The coefficient on the interaction term in the boys' regression is positive in all four groups of subjects (core, classics, science and exact science) but it is not significantly different from zero in all four regressions. This evidence suggest that female teachers' bias have larger effect on boys than male teachers but we do not have enough power to conclusively estimate precisely this difference. The effect of female teachers' biases on girls has the opposite pattern, being smaller than the effect of male teachers' biases in core subjects and in all three high school track subjects. The effect of male teachers on girls remains negative and significant and the effect of female teachers is also negative but smaller and not significantly different than the effect of male teachers. School stereotyping by female teachers seems more harmful for girls but based on our findings

we cannot draw this conclusion with enough statistical conviction. Figure 4 presents the distribution of our main measure of teacher bias for male and female teachers separately.

In Table 11 we allow for the teachers' bias effect to vary by the gender of the school principal. Overall, having a female principal lowers the effect of teachers' biases on boys and increases the effect of these biases on girls. In Table 11 we also allow for an interaction effect between the gender of the teacher and the gender of the principal. The coefficient of the interaction term on the teachers' 11th grade bias is positive in all four groups of subjects and it is statistically significant in the core subjects, and in the science and exact science tracks. This suggests that having a female teacher and principal increases teachers' bias, namely the teacher's bias becomes more in favor of boys. The coefficient of the interaction term is positive in all cases in 12th grade (columns 3 and 4), but it is statistically significant only in the exact science track. Having a female teacher, reduces the teacher bias which means that teachers become more pro-girl. The effect of having a female principal follows a similar pattern for both grades.

Estimated Effect on Dropout and Repetition

In this section we present and discuss estimates of the effect of teachers' grading biases in 11th grade on the dropout rate between 11th and 12th grade and on the repetition of 11th grade. We determine that a student dropped out from schooling at end of 11th grade if he is not enrolled in the same school in 12th grade. The most common reason for a student not to continue schooling in 12th grade in the same high school he attended in earlier grades is a full dropout from schooling. It is rare to find students switching schools between 11th and 12th grade and in most such cases the student will move to a technical vocational schooling which we consider as dropping out from general schooling. In Table 13 we present estimates of teachers' biases in 11th grade on this defined dropout outcomes. We use as treatment the mean teachers' bias in core subjects in 11th grade. The mean dropout rate of boys is 13.3 percent (Table 3) and the effect of teachers' bias is -0.065, significantly different from zero at the 10 percent level of significance. A one standard deviation increase in pro-boys teachers' bias lower the probability of dropping out among boys by 6.5 percent. The estimated effect on girls has the 'wrong' sign but it is not different from zero.

In Table 14 we present estimates of the effect on the likelihood that a student will repeat 11th grade. Repeating 11th grade is a consequence of doing poorly in the national and school exams at the end of 11th grade. The mean repetition rate is 5.3% for boys and 4.5% for girls (Table 3) using the sample of 135 schools. The effect of mean teachers' biases in 11th grade (based on teachers of all subjects in 11th grade) is negative and significant for boys, -0.106 (se=0.046), and not statistically significant for girls (with the 'wrong' sign again). This non-effect on girls is perhaps explained by the lower repetition rate

among girls in our sample. Results based on bias measure based on current own class teachers reveal the same pattern.

5. Effect of Teacher Biases on University Field of Study

The impact of high school teachers' gender biases may affect university enrollment through two channels. The first is by having an effect on test scores on exams that are used for admission to universities and various study programs in higher education institutions. Higher test scores in exams affect the average score and may help students gain an entry to their preferred university and preferred field of study. In addition, higher test scores in the national exams may increase self-confidence and motivation of students, which can increase students' interest in higher education and more challenging and rewarding study programs. In this section we will estimate the effect of teachers' biases on students' choice of field of study conditional on enrollment in a university, and on the quality of the university that a student is enrolled at.

We group field of studies at the university in the same way we do in high school, according to the four study tracks. Humanities includes the university departments of liberal arts, literature, psychology, journalism, philosophy, education, Greek language, history, foreign languages, home economics and law. Social science includes departments of economics, statistics, business and management, accounting, political and European studies. Exact science includes departments of mathematics, engineering, physics and computer science. Science includes departments of biology, chemistry, medicine, pharmacy, veterinary studies and dentistry. From Table 15 we see that among boys, 3.7 percent enrolled in science studies, 22.3 percent in exact science, 21.3 percent in social science and 8.8 in humanities. 18.1 of boys did not enroll in any post-secondary schooling and 25.8 enrolled in vocational schooling. Among girls, 4.9 percent enrolled in science studies, 9.9 percent in exact science, 22.7 percent in social science and 27.7 in humanities. 18.4 of boys did not enroll in any post-secondary schooling and 17.2 enrolled in vocational schooling. Clearly there are large gender differences in the proportion of enrolled students in exact sciences and in humanities. Figure 7 presents the proportion of students enrolled in each field of university study by year and figure 8 presents the proportion of enrolled boys and girls in each field of university study. We therefore focus our analysis on the effect of teachers' high school biases on the choice of field of study conditional on attending university schooling.

We model the choice of students in a linear regression where we stack the four possible choices as the dependent variable for each student against the teachers' bias in each of the four areas of university studies. The dependent variable is a 0/1 indicator, assuming the value of 1 for the observed field of study

and a value 0 for the other three possible choices. We estimate simple linear probability models since a probit or logit models will yield similar estimates given that we use very large samples. We estimate three different specifications: the benchmark includes a year and major fixed effects and the national exam score in 11th or 12th grade, a second specification includes also high school fixed effect and in a third specification we replace the latter with a high school class fixed effect. Standard errors are clustered at the class level.

The teachers' bias that we relate to each possible field of study are as follows. For exact science departments we use the average of the biases in 11^{th} grade in algebra, geometry and physics and the biases in 12^{th} grade in mathematics and physics. For science departments we use the average of the biases in 11^{th} grade in algebra, geometry and physics and the bias in 12^{th} in biology. For humanities departments we use the average of the biases in 11^{th} and 12^{th} grade in history and Modern Greek. For social science we use the average of the 11^{th} grade in Modern Greek and history and the 12^{th} grade bias in economics. Figure 5 and 6 present the 11^{th} and 12^{th} grade average annual teacher bias measured by high school track (figure 5) and the related annual core subjects bias to each field of university study (figure 6).

In Table 16 we present the effect of 11^{th} and 12^{th} grade teachers' biases on the choice of university field of study. The effect on boys is positive and on girls it is negative. The absolute size of the estimated effect of 11^{th} grade bias is similar for boys and girls but for girls the estimates are more precisely measured and they are significantly different from zero. The estimated effect on girls is -0.036 with year and school fixed effects and -0.057 when we add class fixed effects. The estimated effects for girls with school or class fixed effects imply that one standard deviation increase in the bias in favor of boys in a given field lowers the probability of choosing that field of study by 5.7 percent. The respective estimate of the 12^{th} grade bias is similar though less precisely estimated with the class fixed effects. A striking result is that the estimated effect on girls is the same across all three specifications: in column 4 it is -0.035, in column 4 it is -0.035 and in column 6 it is –0.034. The estimates of boys are also relatively stable across specifications. In Table A11 in online appendix we present estimates when the bias is measured based on own class in the current year. The same pattern emerges with a negative effect of 11^{th} and 12^{th} grade bias on girls and a positive though small and insignificant effect on boys.

We next present estimates of the effect of teachers' gender biases on the field of university study based on the school track that students follow in 11^{th} and 12^{th} grade. The outcome variable is a dummy that takes the value of one if a student is enrolled in a university department that is a natural follow-up of their school track. For example, the outcome variable equals to one if students who are in the classics, science and exact science track enroll in humanities, science and exact science departments, respectively. We use the core subjects' related bias as we did before. The estimated effects are overall positive for boys

and negative for girls. A one standard deviation increase in pro-boy teachers' bias in the related core subjects makes girls 14% more likely to study Humanities, given that they follow the classics track in the eleventh grade. The estimated effect on boys is smaller, positive, but it is not different from zero. For students in the science track, we find that a one standard deviation increase in pro-boy teachers' bias in the related core subjects in 11th grade makes boys (girls) 13% (15%) more (less) likely to enroll in a science department, given that are in the science track in the eleventh grade. For students in the exact science track, we find that the effect on girls is negative and statistically significant (-0.13), which for boys it is smaller and insignificant. The estimated effect in the 12th grade is less precisely estimated.

We next present estimates of the effect of teachers' gender biases on the rank of the institution conditional on the student's field of study. We compute two alternative measure ranking measures of universities by field of study. The first is based on the average score in the admission exams of all students by institution and fields of study in 2003, which is the first year in our data. For each field of study, we rank all institution and transform this distribution to percentile rank. The second ranking measure that we use is based on the admission cutoffs in the institution for each field of study. We determine the admission cutoffs by the admission score of the marginal student admitted to the program in 2003. We again use this measure to percentile rank all institutions by field of study.

In Table 19 we present estimates of the effect of teachers' gender biases on the rank of the institution conditional on the student's field of enrollment. We use the gender teacher bias of teachers of subjects that are closely related to the field of study and are part of the core subjects. We find positive estimates for boys and negative for girls. We use as outcome variables both measures of university department ranking that we constructed. A one standard deviation increase in 11th grade pro-boys teachers' bias in the related subjects makes girls enroll in humanities departments that are 12 (column 2) or 10 (column 4) rankings lower in quality. The 12th grade estimates follow the same pattern although they are smaller for girls, yet statistically significant. The estimated effects for boys are positive but insignificant. However, a one standard deviation increase in pro-boy teachers' bias in the related subjects makes boys enroll in science departments that are 28 (column 1) or 22 (column 3) rankings higher in quality. In this case, the estimates for girls are negative but not precisely estimated. We find similar results when we combine related fields of study. For example, girls enroll in lower ranked (7-8 percentile ranks) university departments in humanities or social science, when they have a pro-boy teacher in the related core subjects. Similarly, boys enroll in higher ranked departments in science or exact science when they have a pro-boy teacher in the related core subjects.

In Table 20 we report the effect of 11th and 12th grade teachers' gender bias on the percentile rank of each student institution by field of study, excluding the cohort of 2003. In column 1 we present the estimates for boys using the 11th grade biases as treatment of interest and the institution field of study rank based on the average admission score as the dependent variable. The first row presents estimates when the bias variable is the average in core subjects, the second row it is the bias in the classic track subjects, in the third row it is the bias in science track subjects and in the fourth row it is the bias in exact science track subjects. The outcome variable is the quality of the field of study program at the university, once measured based on the program admission score cutoff (columns 3,4,7,8) and once measured based on mean performance of enrolled students in the program (columns 1,2,5,6). In column 2 we present the respective estimates for girls.

All the estimates in column 1 are positive and all the estimates in column 2 are negative. The effect in the boys sample is statistically significant for the bias in the classic and exact science tracks. For girls the effect is statistically significant for the bias in core subjects and also for the exact science track. The results obtained from the sample of 21 schools and from the sample of 135 schools are similar. The estimates for boys in columns 5 and 7 are all positive and statistically significant (except in exact science) in the sample of 135 schools. A one standard deviation increase in 12th grade pro-boys teachers' bias in the classics and exact science subjects makes girls enroll in departments that are 1-2 rankings (columns 8) lower in quality.

Effect of Teachers' Gender Biases on STEM Field of Studies

Focusing on the effect of teachers' gender biases on students' choice of STEM field of study, we limit the analysis sample to students who are enrolled in high school in the science and exact science tracks. Ninety percent of the students who choose later at the university a STEM field of study are enrolled in one of these two high school tracks. In Panel A of Table 19 we estimate a linear probability choice model (columns 1 and 2) or a Logit (column 3), where we group the field of study options to engineering versus another STEM field of study. In columns 1,2,3 we present the estimates for girls and in columns 4,5,6 we present the estimates for boys. For students in the science track in 12th grade, one standard deviation increase in 12th grade pro-boys teachers' bias in the science track subjects makes boys 0.7%, 0.5% and 1.1% (column 1 ,2 and 3 respectively) more likely to enroll in an engineering department. In panel B we model the choice as mathematics versus another STEM subject. Here, the 11th grade science track bias makes boys more likely to enroll in Mathematics, while the effect on girls is less precise. We find similar results in Panel C, where the choice is between computer sciences versus another STEM subject. In panel D it is engineering, mathematics or computer sciences versus all other STEM

fields of study. One standard deviation increase in 11th grade pro-boy teachers' bias in the science or exact science track subjects makes boys 1.2%, 1.3% and 1.2% (column 1, 2 and 3 respectively) more likely to enroll in an STEM department.

6. Conclusions

In this paper we investigate how teacher gender bias affects students' later academic performance in high school, other education related decisions (drop-out rate and repetition of a grade), the choice of university field of study and the university national rank in this study area in terms of quality of its admitted students. The measure of teachers' gender-biased behavior that we use is based on a comparison between the school classroom boys' and girls' average test scores in a "non-blind" exam that the teacher marks, versus a "blind" exam marked externally. We use panel data information on teachers' class assignment history throughout the period we study and measure the teacher's grading bias in each of his classes. We then use the teacher's average gender bias based on all classes except the current one on which we measure the bias impact. This approach allow to estimate the effect of the persistent component of teachers' biases. Based on observing teachers in seven different classes, we find that the same teachers who are biased for one class are biased in the same way for other classes in the same year and in classes in earlier or later academic years. The very high correlations of within teachers' biases in different classes reveal high persistency in teachers' stereotypical behavior.

For identification, we rely on the random assignments of teachers and students to classes in a large number of high schools in Greece. We use novel data that we collected from a sample of high schools, and we compare students who are exposed to teachers, who might have different patterns of gender stereotypical biases. An important contribution of this paper is the use of gender stereotypical behavior out of sample (other classes) which enable us to address several threats to the interpretation of our findings and demonstrates that our estimates reflect teachers' behavior and not random (small sample) variation in the unobserved "quality" or "non-cognitive" skills of the boys vs. girls in a particular single class or any other class specific dynamics.

Our results may be summarized with three broad conclusions. First, the same teachers who are biased for one class are biased in the same way for other classes in the same year and in classes in earlier or later academic years. The very high correlations of within teachers' biases in different classes reveal high persistency in teachers' stereotypical behavior. Second, teachers' biases in core and elective subjects (classics, social science, science, exact science) have positive effect on boys' and negative effects on girls' performance on end of high school university admission exams. Female teachers are more pro-girls

on average but the effect of female and male teachers' biases on national exams are not statistically different. Third, teachers' biases in specific courses lower the likelihood that students enroll in a related field of study at the university. This average effect masks large heterogeneity by gender, being larger and statistically significant for girls and not different from zero for boys. Focusing on the choice of STEM field of study among students enrolled in the science and exact science tracks in high school, we find that teachers' gender biases impact mainly boys choices of engineering, mathematics and computer engineering.

7. References

Alesina, A., P. Giuliano, and N. Nunn. 2013. "On the Origin of Gender Roles: Women and the Plough", *Quarterly Journal of Economics* 128(2), 469-530.

Adams, B.N., 1972. "Birth Order: A Critical Review", *Sociometry*, 35(3), 411-439.

Bae, Y. and T.M. Smith, 1997. "Women in Mathematics and Science". Findings from "The Condition of Education", *National Center for* Education Statistics 1997, no. 11.

Becker, G.S., W.H. Hubbard and K.M. Murphy, 2010. "Explaining the Worldwide Boom in Higher Education of Women", *Journal of Human Capital* 4, 203-241.

Bertrand, M., D. Chugh and S. Mullainathan, 2005, "Implicit Discrimination", *American Economic Review*, 95(2), 94-98.

Bertrand, M. and E. Duflo, "Field Experiments on Discrimination", Forthcoming in: A. Banerjee and E. Duflo (Eds.), Handbook of Field Experiments.

Bian, L. Leslie, S.J, & Cimpian, A. (2017). Gender stereotypes About Intellectual Ability Emerge Early and Influence Children's Interests. *Science, 355(6323),* 389-391.

Björn, T.H., Höglin, E. and M. Johannesson, 2011. "Are Boys Discriminated in Swedish High Schools?", *Economics of Education Review* 30(4), 682-690.

Blank, R.M., 1991. "The Effects of Double-Blind versus Single-Blind Reviewing: Experimental Evidence from the American Economic Review", *American Economic Review* 81, 1041-1067.

Botelho, F., Mdeira, R.A. and M.A., Rangel 2015. "Racial Discrimination in Grading: Evidence from Brazil", *American Economic Journal: Applied Economics*, 7(4), 37-52.

Burgess, S. and E. Greaves, 2013. "Test Scores, Subjective Assessment, and Stereotyping of Ethnic Minorities", J*ournal of Labor Economics* 31, 535-576.

Burnsztyn, L. and R. Jensen, 2015, "How Does Peer Pressure Affect Educational Investments?", The *Quarterly Journal of Economics*, 130(3), 1329-1367.

Dee, T. S., 2005. "A Teacher Like Me: Does Race, Ethnicity, or Gender Matter?" *American Economic Review*, 95(2), 158-165.

Ellison, G. and A. Swanson, 2010. "The Gender Gap in Secondary School Mathematics at High Achievement Levels: Evidence from the American Mathematics Competitions", *Journal of Economic Perspectives* 24, 109-128.

Fershtman, C. and U. Gneezy, 2001. "Discrimination in a Segmented Society: An Experimental Approach", *The Quarterly Journal of Economic* 116(1), 351-377.

Goldin, C., "A Grand Gender Convergence: Its Last Chapter", *American Economic Review*. Forthcoming 104.

Goldin, C., L.F. Katz and I. Kuziemko, 2006. "The Homecoming of American College Women: The Reversal of the College Gender Gap", Journal of Economic Perspectives 20, 133-156.

Goldin, C. and C. Rouse, 2000. "Orchestrating Impartiality: The Impact of "Blind "Auditions on Female Musicians", *The American Economic Review* 90, 715-741.

Goulas, S. and R. Megalokonomou, 2016 "Which degrees do students prefer during recessions?", MPRA Working Paper No. 75355

Guiso, L., F. Monte, P. Sapienza and L. Zingales, 2008. "Culture, Gender, and Math", *Science* 320, 1164-1165.

Hanna, R.N., and L.L., Linden, 2012. "Discrimination in Grading", *American Economic Journal: Economic Policy*, 4(4), 146-68.

Lavy, V., 2008. "Do Gender Stereotypes Reduce Girls' or Boys' Human Capital Outcomes? Evidence from a Natural Experiment", *Journal of Public Economics* 92, 2083-2105.

Lavy, V., 2011. "What Makes an Effective Teacher? Quasi-Experimental Evidence", Forthcoming CESifo Economic Studies.

Lavy, V. and E. Sand, "On the Origins of the Gender Human Capital Gap: Short and Long Term Effect of Teachers' Stereotypes", Draft, Applied Micro Seminar, Department of Economics, Hebrew University of Jerusalem, July 2014.

Leslie, S.J., A. Cimpian, M. Meyer and E. Freeland, 2015. "Expectations of Brilliance underlie Gender Distributions across Academic Disciplines", Science 347, 262-26.

Machin, S. and T. Pekkarinen, 2008. "Global Sex Differences in Test Score Variability", *Science* 322, 1331-1332.

Niederle, M. and L. Vesterlund, 2007. "Do Women Shy Away from Competition? Do Men Compete Too Much?" *The Quarterly Journal of Economics* 122, 1067-1101.

Paglin, M. and A.M. Rufolo, 1990. "Heterogeneous Human Capital, Occupational Choice, and Male-Female Earnings Differences", *Journal of Labor Economics* 8, 123-144.

Reuben, E., Sapienza P. and L. Zingales, 2014. 'How Stereotypes Impair Women's Careers in Science," *Proceeding of the National Academy of Science*, Forthcoming.

Spolaore, E. and R. Wacziarg, 2009, "The Diffusion of Development", *The Quarterly Journal of Economics* 124, 469-529.

Terrier, C., 2014, "Giving a Little Help to Girls? Evidence on Grade Discrimination and its Effect on Students' Achievement", PSE Working Papers n. 2014-36.

Table 1: Descriptive statistics for the sample of 135 schools

Variable	Mean	Std. Dev.	Min.	Max.
Student Characteristics				
11th grade				
Female	0.563	0.497	0	1
Total absences (in hours per year)	50.938	27.506	1	450
Proportion of students by track:				
Classics	0.367	0.482	0	1
Exact Sciences	0.279	0.449	0	1
Exact Science	0.344	0.475	0	1
GPA 11	72.321	14.156	0	100
Repeat 11th grade	0.023	0.149	0	1
12th grade				
Female	0.562	0.496	0	1
Total absences (in hours per year)	73.444	30.785	1	208
Proportion of students by track:				
Classics	0.370	0.438	0	1
Sciences	0.159	0.366	0	1
Exact Science	0.463	0.498	0	1
GPA 12	76.976	12.531	44	100
Age	17.902	0.465	15	54
School Characteristics				
Private School	0.037	0.190	0	1
Experimental School	0.044	0.207	0	1
Public School	0.919	0.274	0	1
Urban	0.896	0.306	0	1
Postcode Income(in 2009 Euro)	22,450	7,945	11,784	66,521
University Enrollment Characteristics				
University admission national exam score	64.987	20.178	10.35	99.3
Retake the national exams	0.113	0.317	0	1
Number of reported options in preference list	25.151	22.295	1	257
Rank of enrolled option in preference list	8.397	10.617	1	242
Enrollment in university or vocational schooling	0.817	0.387	0	1
Enrol in Exact Science department	0.153	0.360	0	1
Enrol in Science department	0.042	0.201	0	1
Enrol in Humanities department	0.193	0.407	0	1
Enrol in Social Science department	0.220	0.413	0	1
Enrol in vocational schooling	0.209	0.407	0	1

Note: The variable "Repeat 11th grade" indicates repetition due to poor performance in the school and national exams. The variables "GPA11" and "GPA12" include the average over the school exam scores in the first and second term, in 11th and 12th grade, respectively. Total absences are measured in hours per year.

Table 2: Descriptive Statistics by study sample

11th grade sample 2003-2005 and 12th grade sample 2003-2011						
	116 schools		21 schools		Difference	
Variable	**Mean**	**(sd)**	**Mean**	**(sd)**	**diff**	**(s.e.)**
11th grade						
Number of classes	3.900	(1.134)	3.923	(1.581)	-0.023	(0.014)
Class size	19.537	(5.073)	18.653	(4.818)	0.884	(0.059)
School cohort size	75.868	(26.257)	75.669	(34.945)	0.198	(0.320)
Proportion of students by track						
Classics	0.365	(0.058)	0.375	(0.054)	-0.010	(0.001)
Science	0.282	(0.070)	0.265	(0.095)	0.017	(0.001)
Exact Science	0.342	(0.068)	0.360	(0.078)	0.018	(0.001)
Proportion of female students	0.562	(0.496)	0.573	(0.495)	-0.011	(0.007)
Teachers with a bias measure in year t in own class			1,346			
Teachers with a bias measure in year t in other classes			1,158			
Teachers with a bias measure biases in year t-i			764			
Teachers with a bias measure in year t+i			882			
Teachers with a bias measure in all other classes in any year			1,289			
12th grade						
Number of classes	3.868	(1.143)	3.854	(1.546)	0.014	(0.018)
Class size	19.675	(4.959)	19.006	(4.924)	0.669	(0.075)
School cohort size	75.880	(26.252)	75.667	(34.943)	0.213	(0.461)
Proportion of students by track						
Classics	0.368	(0.060)	0.376	(0.056)	-0.008	(0.007)
Science	0.159	(0.049)	0.164	(0.056)	-0.005	(0.001)
Exact Science	0.463	(0.071)	0.460	(0.064)	0.003	(0.001)
Proportion of female students	0.562	(0.496)	0.573	(0.495)	-0.011	(0.007)
Age	17.903	(0.451)	17.892	(0.552)	0.011	(0.006)
Number of teacher biases in year t in own class			2,916			
Teachers with a bias measure in year t in other classes			1,839			
Teachers with a bias measure in year t-i			1,620			
Teachers with a bias measure in year t+i			1,724			
Teachers with a bias measure in all other classes in any year			2,536			

Notes: There are three tracks available to students in 11th and 12th grade: classics, science and exact science. In 11th grade the subjects taught in the classics track are ancient Greek, philosophy and latin; in the science track: mathematics, physics, chemistry; and in the exact science track: mathematics, physics and computer science. In 12th grade the subjects taught in the classics track are ancient Greek, latin, literature and history; in science track: biology, mathematics, physics and chemistry; and in exact science track: mathematics, physics, business administration and application development.

Table 3: Descriptive Statistics by Gender

Dependent Variable: Dummy for drop out and Transfers			
	(1)	(2)	(3)
Variable	Males	Females	(1)-(2)
	(sd)	(sd)	(se)
Repetition Rate of 11th grade	0.053	0.045	0.009
	(0.006)	(0.005)	(0.008)
Total Absences in 11th grade (in hours per year)	52.538	53.869	-1.331
	(0.741)	(0.703)	(1.030)
Drop out rate between 11th and 12th grade	0.133	0.112	0.021
	(0.002)	(0.001)	(0.002)

Note: A student repeats the 11th grade when his academic performance in school and national exams is poor. The repetition rate is measured using the sample of 135 schools. The drop out rate and the total absences in 11th grade are measured using the sample of 21 schools.

Table 4: Mean Scores and Standard Deviations in the National Exam (blind) and the School Exam (non-blind) in 11th Grade, 2003-2005, Sample of 21 Schools

	National Exam			School Exam		
Variables	Boys	Girls	**Difference**	Boys	Girls	**Difference**
	(sd)	(sd)	(se)	(sd)	(sd)	(se)
	(1)	(2)	(1)-(2)	(3)	(4)	(3)-(4)
Core subjects						
Modern Greek	-0.433	-0.039	-0.394	-0.179	0.322	-0.501
	(1.051)	(1.014)	(0.037)	(1.023)	(0.925)	(0.035)
History	-0.247	-0.041	-0.206	-0.112	0.230	-0.342
	(0.942)	(1.003)	(0.035)	(1.018)	(0.958)	(0.035)
Physics	-0.033	-0.093	0.059	0.032	0.070	-0.038
	(0.993)	(1.021)	(0.036)	(0.961)	(0.921)	(0.033)
Algebra	-0.059	-0.067	0.008	-0.018	0.073	-0.090
	(1.011)	(1.015)	(0.036)	(1.006)	(0.955)	(0.034)
Geometry	-0.045	-0.076	0.031	-0.013	0.091	-0.104
	(0.993)	(0.966)	(0.035)	(1.014)	(0.986)	(0.036)
Track: Classics						
Ancient Greek	-0.347	-0.047	-0.300	-0.188	0.176	-0.364
	(0.996)	(0.997)	(0.075)	(1.024)	(0.978)	(0.074)
Philosophy	-0.378	-0.092	-0.286	-0.146	0.109	-0.254
	(0.918)	(1.003)	(0.074)	(1.020)	(0.917)	(0.070)
Latin	-0.338	-0.022	-0.316	-0.110	0.231	-0.343
	(1.044)	(1.009)	(0.076)	(0.973)	(0.813)	(0.063)
Track: Science						
Mathematics	-0.146	-0.052	-0.094	-0.079	0.090	-0.169
	(1.076)	(1.012)	(0.069)	(1.044)	(1.012)	(0.070)
Physics	-0.130	-0.038	-0.091	0.025	0.158	-0.133
	(1.122)	(1.007)	(0.072)	(0.933)	(0.868)	(0.069)
Chemistry	-0.095	-0.026	-0.069	-0.147	-0.026	-0.122
	(1.036)	(0.992)	(0.069)	(1.033)	(0.991)	(0.061)
Track: Exact Science						
Mathematics	-0.160	0.105	-0.264	-0.090	0.282	-0.372
	(0.987)	(1.051)	(0.060)	(0.998)	(0.924)	(0.058)
Physics	-0.130	0.062	-0.192	-0.008	0.220	-0.228
	(0.982)	(1.045)	(0.060)	(0.927)	(0.909)	(0.055)
Computer Science	-0.084	0.010	-0.094	-0.083	0.275	-0.358
	(0.994)	(1.076)	(0.062)	(0.960)	(0.857)	(0.055)

Notes: The national and school exam scores are standardized z-scores. This table presents test scores gender gaps by type of exam (blind and non-blind) and subject in 11th grade. A positive difference means that boys outperform girls, while a negative difference means that girls outperform boys. The non-blind score in each subject is the score in the second term school exam.

Table 5: Mean Scores and Standard Deviations in the National Exam (blind) and the School Exam (non-blind) in 12th Grade, 2003-2011, Sample of 21 Schools

	National Exam			School Exam		
Variables	Boys (sd)	Girls (sd)	**Difference** (se)	Boys (sd)	Girls (sd)	**Difference** (se)
	(1)	(2)	(1)-(2)	(3)	(4)	(3)-(4)
Modern Greek	-0.392	0.015	-0.395	-0.177	0.301	-0.478
	(1.023)	(1.005)	(0.029)	(1.053)	(0.885)	(0.028)
History	-0.250	-0.070	-0.180	-0.089	0.223	-0.312
	(0.981)	(1.017)	(0.039)	(0.978)	(0.892)	(0.036)
Mathematics	-0.084	-0.100	0.017	-0.017	0.120	-0.136
	(0.998)	(1.009)	(0.032)	(0.922)	(1.052)	(0.031)
Physics	-0.015	-0.099	0.085	0.093	0.185	-0.092
	(0.977)	(1.006)	(0.038)	(0.874)	(0.852)	(0.033)
Biology	-0.194	-0.118	-0.076	-0.015	0.086	-0.239
	(1.000)	(1.023)	(0.035)	(1.104)	(0.898)	(0.036)
Track: Classics						
Ancient Greek	-0.357	-0.030	-0.327	0.003	0.279	-0.276
	(0.990)	(0.971)	(0.058)	(0.991)	(0.856)	(0.051)
Latin	-0.313	0.047	-0.360	-0.154	0.210	-0.364
	(0.989)	(0.946)	(0.056)	(1.041)	(0.850)	(0.052)
Modern Literature	-0.408	-0.057	-0.351	-0.163	0.148	-0.310
	(1.038)	(1.026)	(0.061)	(1.040)	(0.889)	(0.054)
History	-0.212	-0.114	-0.096	-0.054	0.196	-0.250
	(0.994)	(1.014)	(0.059)	(1.031)	(0.908)	(0.053)
Track: Science						
Biology	-0.189	-0.095	-0.093	-0.058	-0.003	-0.055
	(1.076)	(1.029)	(0.073)	(0.979)	(0.874)	(0.071)

Continued on next page

Table 5 – *Continued from previous page*

	National Exam			School Exam		
Variables	Boys (sd) (1)	Girls (sd) (2)	**Difference** (se) (1)-(2)	Boys (sd) (3)	Girls (sd) (4)	**Difference** (se) (3)-(4)
Mathematics	-0.046 (1.024)	-0.106 (0.962)	0.060 (0.57)	-0.013 (1.096)	0.119 (0.971)	-0.132 (0.059)
Physics	0.031 (0.992)	-0.128 (0.962)	0.159 (0.073)	0.063 (0.928)	0.136 (0.805)	-0.073 (0.061)
Chemistry	0.004 (0.987)	-0.046 (0.987)	0.050 (0.074)	-0.108 (1.080)	-0.017 (0.941)	-0.091 (0.074)
Track: Exact Science						
Mathematics	-0.151 (0.967)	0.024 (0.952)	-0.175 (0.041)	-0.111 (1.019)	0.221 (0.936)	-0.332 (0.042)
Physics	-0.116 (0.977)	-0.030 (0.973)	-0.087 (0.041)	0.057 (0.901)	0.302 (0.830)	-0.246 (0.037)
Business Administration	-0.211 (0.982)	-0.040 (1.022)	-0.171 (0.042)	-0.112 (1.015)	0.283 (0.784)	-0.395 (0.039)
Computer Science	-0.071 (0.999)	-0.009 (0.982)	-0.062 (0.038)	-0.112 (1.038)	0.188 (0.888)	-0.301 (0.042)
Optional						
Economics	-0.136 (0.982)	-0.024 (1.009)	-0.112 (0.038)	-0.003 (0.933)	0.203 (0.832)	-0.206 (0.034)

Note: The national and school exam scores are standardized z-scores. This table presents test scores gender gaps by type of exam (blind and non-blind) and subject in 12th grade. A positive difference means that boys outperform girls, while a negative difference means that girls outperform boys. The non-blind score in each subject is the score in the second term school exam.

Table 6: Descriptive Statistics for 11th and 12th Grade Teachers, Sample of 21 Schools

Variable	Mean	Std. Dev.	Min.	Max.
11th grade				
Number of classes taught by a teacher	7.25	4.82	1	23
Number of classes/subjects taught by teacher by year	3.40	2.04	1	10
2003	3.50	2.07	1	10
2004	3.22	2.00	1	10
2005	3.49	2.03	1	9
Number of different subjects taught by teacher by year	1.62	0.63	1	3
2003	1.67	0.67	1	3
2004	1.59	0.62	1	3
2005	1.59	0.62	1	3
Number of different classes taught by teacher by year	1.74	0.89	1	4
2003	1.75	0.85	1	4
2004	1.63	0.88	1	4
2005	1.84	0.93	1	4
Number of years a teacher teaches by year	2.18	0.87	1	3
12th grade				
Number of classes taught by teacher	10.76	8.77	1	46
Number of classes/subjects taught by teacher by year	2.65	1.60	1	9
2003	2.98	1.74	1	7
2004	3.00	1.70	1	7
2005	3.05	1.81	1	8
2006	2.71	1.83	1	9
2007	2.25	1.22	1	6
2008	2.17	1.13	1	5
2009	2.28	1.31	1	6
2010	2.22	1.16	1	5
2011	2.43	1.45	1	6
Number of different subjects taught by teacher by year	1.43	0.67	1	4
2003	1.53	0.74	1	4
2004	1.54	0.71	1	4
2005	1.56	0.64	1	4
2006	1.49	0.83	1	4
2007	1.27	0.48	1	3
2008	1.33	0.50	1	3
2009	1.33	0.58	1	3
2010	1.33	0.62	1	3
2011	1.40	0.73	1	4
Number of different classes taught by teacher by year	1.51	0.08	1	4
2003	1.60	0.76	1	4
2004	1.66	0.92	1	4
2005	1.80	1.04	1	4
2006	1.45	0.68	1	4
2007	1.39	0.62	1	3
2008	1.34	0.58	1	3
2009	1.32	0.54	1	3
2010	1.31	0.60	1	3
2011	1.36	0.77	1	4
Number of years a teacher teaches	4.39	2.34	1	9

Notes: The sample includes all teachers who teach core or track subjects in 11th and 12th grade.

Table 7: Descriptive Statistics for Different Measures of Teacher Bias in 11th and 12th Grade, Sample of 21 Schools

	Prop. of Fem. Teachers	Teacher Bias measured in other classes (21 schools)		Teacher Bias measured in the own class (21 schools)				**correlation between**
Variable		**Mean**	**(sd)**	**Mean**	**(sd)**	**Diff.**	**se**	(2)and(4)
Bias in	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
11th grade (2003-2005)								
Core subjects								
Modern Greek	0.71	-0.100	(0.369)	-0.097	(0.463)	-0.003	(0.004)	0.85
History	0.67	-0.110	(0.307)	-0.129	(0.378)	0.019	(0.004)	0.80
Algebra	0.39	-0.095	(0.236)	-0.106	(0.306)	0.010	(0.003)	0.78
Geometry	0.37	-0.102	(0.253)	-0.094	(0.303)	-0.008	(0.003)	0.78
Physics	0.45	-0.104	(0.269)	-0.094	(0.319)	-0.010	(0.003)	0.81
Classics Track								
Ancient Greek	0.63	-0.160	(0.355)	-0.152	(0.396)	-0.008	(0.007)	0.80
Philosophy	0.66	-0.061	(0.363)	-0.027	(0.415)	-0.034	(0.007)	0.79
Latin	0.69	-0.116	(0.284)	-0.087	(0.372)	-0.029	(0.007)	0.78
Science Track								
Mathematics	0.46	-0.074	(0.224)	-0.066	(0.326)	-0.008	(0.008)	0.72
Physics	0.41	-0.004	(0.246)	-0.018	(0.332)	0.013	(0.007)	0.75
Chemistry	0.37	-0.095	(0.323)	-0.077	(0.351)	-0.018	(0.006)	0.87
Exact Science Track								
Mathematics	0.32	-0.057	(0.258)	-0.080	(0.307)	0.024	(0.010)	0.86
Physics	0.39	-0.104	(0.243)	-0.105	(0.334)	0.0002	(0.005)	0.84
Technology and Computers	0.29	-0.223	(0.338)	-0.248	(0.397)	0.025	(0.007)	0.82
12th grade (2003-2011)								
Core subjects								
Modern Greek	0.59	-0.045	(0.365)	-0.062	(0.500)	0.017	(0.004)	0.76
Biology	0.20	-0.112	(0.429)	-0.158	(0.667)	0.046	(0.007)	0.60
History	0.48	-0.113	(0.319)	-0.157	(0.409)	0.044	(0.005)	0.67
Mathematics	0.29	-0.128	(0.336)	-0.128	(0.512)	0.0003	(0.004)	0.72
Physics	0.020	-0.172	(0.276)	-0.193	(0.328)	0.021	(0.004)	0.69
Classics Track								
Ancient Greek	0.53	-0.060	(0.341)	-0.039	(0.396)	-0.021	(0.003)	0.81
Latin	0.64	-0.101	(0.300)	-0.080	(0.388)	-0.021	(0.003)	0.77
Literature	0.57	-0.108	(0.352)	-0.069	(0.502)	-0.040	(0.004)	0.80
History	0.58	-0.150	(0.292)	-0.178	(0.370)	0.029	(0.003)	0.75
Science Track								
Biology	0.25	-0.141	(0.373)	-0.089	(0.587)	-0.052	(0.005)	0.73
Mathematics	0.13	-0.195	(0.381)	-0.203	(0.511)	0.008	(0.005)	0.68
Physics	0.20	-0.231	(0.283)	-0.268	(0.462)	0.037	(0.004)	0.74
Chemistry	0.19	-0.169	(0.408)	-0.147	(0.529)	-0.022	(0.005)	0.71
Exact Science Track								
Mathematics	0.27	-0.126	(0.284)	-0.138	(0.327)	0.012	(0.003)	0.76
Physics	0.21	-0.193	(0.279)	-0.184	(0.345)	-0.009	(0.003)	0.69
Business Administration	0.58	-0.134	(0.313)	-0.150	(0.401)	0.016	(0.003)	0.73
Computers	0.35	-0.182	(0.273)	-0.191	(0.373)	0.008	(0.003)	0.68
Optional								
Economics	0.56	-0.108	(0.307)	-0.065	(0.440)	-0.044	(0.004)	0.76

Notes: Negative bias means that the teacher is pro-girl. The means are weighted by number of students.

Table 8: Correlations Between Different Measures of Teacher Bias

Dependent Variable: Current year own teacher bias				
	11th grade		12th grade	
	(1)	(2)	(3)	(4)
	Panel A: All Teachers			
Bias measured in other classes in same year	0.813	0.801	0.612	0.606
	(0.035)***	(0.037)***	(0.037)***	(0.038)***
Sample Size	818	818	1,279	1,279
Bias measured in other classes in any year	0.720	0.704	0.731	0.723
	(0.052)***	(0.049)***	(0.033)***	(0.033)***
Sample Size	844	844	1,895	1,895
	Panel B: Female Teachers			
Bias measured in other classes in same year	0.771	0.756	0.620	0.510
	(0.054)***	(0.061)***	(0.061)***	(0.070)***
Sample Size	414	414	501	501
Bias measured in other classes in any year	0.708	0.705	0.715	0.638
	(0.085)***	(0.081)***	(0.055)***	(0.064)***
Sample Size	426	426	761	761
	Panel C: Male Teachers			
Bias measured in other classes in same year	0.867	0.820	0.597	0.562
	(0.036)***	(0.038)***	(0.047)***	(0.049)***
Sample Size	404	404	778	778
Bias measured in other classes in any year	0.731	0.628	0.737	0.710
	(0.058)***	(0.060)***	(0.042)***	(0.043)***
Sample Size	423	423	1,134	1,134
Subjects FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
School FE		✓		✓

Notes: The sample includes all teachers who teach core and track subjects. Standard errors are clustered by school and are reported in parentheses. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

Table 9: Effect of 11th Grade Teacher Bias Measured in All Other Classes On Blind Score in 12th Grade, Sample of 21 Schools

Dependent Variable: Blind score in 12th grade national exams						
	BOYS			GIRLS		
	(1)	(2)	(3)	(4)	(5)	(6)
Core Subjects						
	0.047	0.065	0.103	-0.100	-0.085	-0.111
	(0.040)	(0.034)*	(0.037)***	(0.041)**	(0.039)**	(0.043)***
Sample Size	9,406	9,406	9,406	11,844	11,844	11,844
Classics Subjects						
	0.067	0.088	0.067	-0.086	-0.063	-0.067
	(0.089)	(0.064)	(0.063)	(0.046)*	(0.042)	(0.046)
Sample Size	1,817	1,817	1,817	7,080	7,080	7,080
Science Subjects						
	0.036	0.059	0.128	-0.094	-0.082	-0.088
	(0.069)	(0.056)	(0.066)*	(0.047)**	(0.045)*	(0.047)*
Sample Size	3,236	3,236	3,236	3,386	3,386	3,386
Exact Science Subjects						
	0.012	0.009	0.051	-0.107	-0.102	-0.154
	(0.048)	(0.043)	(0.046)	(0.066)	(0.061)*	(0.067)**
Sample Size	5,703	5,703	5,703	3,364	3,364	3,364
Subjects FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
School FE		✓	✓			✓
Class FE			✓			✓

Notes: The datasets for the core subjects and each track subjects include stacked observations for each subject/exam. Each row presents estimates from separate OLS regressions. All specifications include the students' blind score as a control. The second panel "Classics Subjects" includes relevant exams from the core and the classics track. The third panel "Science Subjects" includes relevant exams from the core and the science track. The forth panel "Exact Science Subjects" includes relevant exams from the core and the exact science track. Standard errors are clustered by class and are reported in parentheses. All scores are standardised z-scores. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

Table 10: Heterogeneity in the Effect of 11th Grade Gender Bias (measured in all other classes) on Blind 12th Grade Score by the Gender of the Teacher

Dependent Variable: Blind score in 12th grade national exams						
	BOYS			GIRLS		
	(1)	(2)	(3)	(4)	(5)	(6)
Core Subjects						
Bias	0.033	0.031	0.066	-0.073	-0.087	-0.184
	(0.055)	(0.050)	(0.048)	(0.052)	(0.051)*	(0.057)***
Bias *Female Teacher	0.030	0.046	0.040	-0.061	-0.012	0.117
	(0.080)	(0.070)	(0.072)	(0.084)	(0.082)	(0.087)
Female Teacher	0.014	-0.012	-0.011	-0.028	-0.041	-0.034
	(0.026)	(0.023)	(0.023)	(0.027)	(0.026)	(0.028)
Sample Size	9,609	9,609	9,609	12,082	12,082	12,082
Classics Subjects						
Bias	0.038	0.040	0.057	-0.055	-0.062	-0.147
	(0.054)	(0.048)	(0.046)	(0.049)	(0.049)	(0.054)***
Bias *Female Teacher	0.019	0.039	0.049	-0.083	-0.033	0.081
	(0.078)	(0.067)	(0.069)	(0.077)	(0.074)	(0.078)
Female Teacher	0.016	-0.007	-0.006	-0.020	-0.028	-0.017
	(0.026)	(0.023)	(0.023)	(0.025)	(0.023)	(0.026)
Sample Size	9,961	9,961	9,961	13,337	13,337	13,337
Science Subjects						
Bias in 11th grade	0.025	0.018	0.050	-0.075	-0.097	-0.197
	(0.054)	(0.050)	(0.048)	(0.051)	(0.050)*	(0.056)***
Bias *Female Teacher	0.032	0.058	0.056	-0.072	-0.017	0.112
	(0.078)	(0.069)	(0.072)	(0.083)	(0.081)	(0.085)
Female Teacher	0.016	-0.011	-0.010	-0.028	-0.044	-0.038
	(0.025)	(0.023)	(0.023)	(0.027)	(0.025)*	(0.028)
Sample Size	9,956	9,956	9,956	12,433	12,433	12,433
Exact Science Subjects						
Bias in 11th grade	0.035	0.025	0.061	-0.067	-0.088	-0.186
	(0.052)	(0.047)	(0.048)	(0.050)	(0.051)*	(0.057)***
Bias *Female Teacher	0.021	0.052	0.042	-0.066	-0.006	0.122
	(0.077)	(0.068)	(0.072)	(0.084)	(0.081)	(0.087)
Female Teacher	0.015	-0.010	-0.015	-0.029	-0.041	-0.036
	(0.025)	(0.023)	(0.023)	(0.027)	(0.025)	(0.027)
Sample Size	10,260	10,260	10,260	12,462	12,462	12,462
Subjects FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
School FE		✓	✓			✓
Class FE			✓			✓

Notes: Standard errors are clustered by class and are reported in parentheses. The second panel "classics subjects" includes all relevant exams from the core and the classics track. The third panel "science subjects" includes all relevant exams from the core and the science track. The forth panel "exact science subjects" includes all relevant exams from the core and the exact science track. The scores are standardised z-scores. All specifications include the students' blind score as a control. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

Table 11: Heterogeneity in the Effect of 11th Grade Gender Bias (measured in all other classes) on Blind 12th Grade Score by the Gender of the Principal

Dependent Variable: Blind score in 12th grade national exams		
	BOYS	GIRLS
	(1)	(2)
Core Subjects		
Bias	0.081	-0.062
	(0.051)	(0.051)
Bias *Female Principal	-0.183	-0.058
	(0.098)*	(0.093)
Female Principal	-0.098	-0.047
	(0.046)**	(0.038)
Sample Size	6,794	8,924
Classics Subjects		
Bias	0.065	-0.041
	(0.115)	(0.061)
Bias*Female Principal	0.013	-0.057
	(0.161)	(0.091)
Female Principal	0.060	-0.014
	(0.112)	(0.049)
Sample Size	1,459	5,414
Science Subjects		
Bias	0.077	-0.078
	(0.074)	(0.059)
Bias *Female Principal	-0.409	0.142
	(0.122)***	(0.133)
Female Principal	-0.266	-0.083
	(0.068)***	(0.068)
Sample Size	2,249	2,396
Exact Science Subjects		
Bias	0.080	-0.046
	(0.059)	(0.070)
Bias*Female Principal	-0.173	-0.253
	(0.122)	(0.178)
Female Principal	-0.087	-0.031
	(0.053)	(0.054)
Sample Size	4,146	2,678
Subjects FE	✓	✓
Year FE	✓	✓

Notes: Standard errors are clustered by class and are reported in parentheses. The second panel "classics subjects" includes all relevant exams from the core and the classics track. The third panel "science subjects" includes all relevant exams from the core and the science track. The forth panel "exact science subjects" includes all relevant exams from the core and the exact science track. The scores are standardised z-scores. All specifications include the students' blind score as a control. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

Table 12: Heterogeneity in Teacher Bias (measured in all other classes) by the Gender of the Teacher and the Principal

Dependent Variable: Teacher Bias in all other classes				
	11th Grade		12th Grade	
	(1)	(2)	(3)	(4)
Core Subjects				
Female Teacher * Female Principal	0.202	0.228	0.101	0.109
	(0.099)**	(0.103)**	(0.077)	(0.080)
Female Teacher	-0.060	-0.080	0.014	-0.001
	(0.038)	(0.038)**	(0.048)	(0.044)
Female Principal	-0.065		-0.044	
	(0.067)		(0.083)	
Sample Size	660	660	292	320
Classics Subjects				
Female Teacher * Female Principal	0.144	0.141	0.049	0.034
	(0.128)	(0.121)	(0.070)	(0.071)
Female Teacher	-0.056	-0.061	-0.066	-0.062
	(0.040)	(0.046)	(0.039)*	(0.043)
Female Principal	-0.070		-0.230	
	(0.136)		(0.112)*	
Sample Size	901	901	676	676
Science Subjects				
Female Teacher * Female Principal	0.192	0.248	0.045	0.068
	(0.163)	(0.147)*	(0.084)	(0.081)
Female Teacher	-0.067	-0.079	-0.048	-0.063
	(0.037)*	(0.042)*	(0.040)	(0.041)
Female Principal	-0.010		-0.061	
	(0.134)	(0.074)		
Sample Size	842	842	598	598
Exact Science Subjects				
Female Teacher * Female Principal	0.236	0.262	0.137	0.128
	(0.104)*	(0.108)**	(0.071)*	(0.072)*
Female Teacher	-0.092	-0.133	-0.064	-0.051
	(0.038)**	(0.039)***	(0.033)*	(0.035)
Female Principal	-0.109		-0.114	
	(0.059)*		(0.048)**	
Sample Size	837	837	876	876
Subjects FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
School FE		✓		✓

Notes: Standard errors are clustered by class and are reported in parentheses. The second panel "classics subjects" includes all relevant exams from the core and the classics track. The third panel "science subjects" includes all relevant exams from the core and the science track. The forth panel "exact science subjects" includes all relevant exams from the core and the exact science track. The scores are standardised z-scores. A teacher sample is used. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

Table 13: The Effect of Teacher bias on Drop Out Rate Between 11th and 12th Grade

Dependent Variable: Drop-out rate between 11th and 12th grades				
11th grade sample 2003-2005				
	Boys		**Girls**	
	(1)	(2)	(3)	(4)
Teacher Bias in 11th grade in all subjects	-0.071	-0.065	-0.024	-0.044
	(0.032)**	(0.035)*	(0.036)	(0.035)
Sample Size	10,410	10,410	12,753	12,753
Year FE	✓	✓	✓	✓
Subjects FE	✓	✓	✓	✓
School FE		✓		✓

Notes: All specifications include a constant. A sample of 135 schools is used. We include total absences in 11th grade and students' blind score in 11th grade as controls. A student drops out between 11th and 12th grades when he/she appears in our data in 11th grade but not in 12th grade. If a student drops out between 11th and 12th grades, it implies either that he drops out from school or switches school. The latter is more likely to happen if a student decides to pursue a different type of school studies (vocational studies). Standard errors are clustered by school and are reported in parentheses. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

Table 14: The Effect of Teacher Bias on Students' Repetition of 11th Grade

Dependent Variable: Retake the Grade				
11th grade sample 2003-2005				
	Boys		**Girls**	
	(1)	(2)	(3)	(4)
	11th grade			
	Sample of 21 schools			
Other Classes Teacher Bias	-0.077	-0.106	-0.041	-0.040
	(0.039)*	(0.046)**	(0.034)	(0.039)
Sample Size	1,397	1,397	1,734	1,734
	Sample of 21 schools			
Own Current Teacher Bias in all subjects	-0.051	-0.054	-0.008	0.007
	(0.027)*	(0.029)*	(0.039)	(0.047)
Sample Size	1,648	1,648	2,065	2,065
Year FE	✓	✓	✓	✓
School FE		✓		✓

A student has to retake a grade when he fails in the end-of-year exams. The "own teacher bias in all subjects" is the overall bias a student is exposed in 11th grade. In all specifications we include students' blind score and total absences in 11th grade (hours per year) as a control. The "other classes teacher bias in all subjects" is the overall bias teacher exhibited in all other classes in 11th grade. Failing the end-of-year exams means that a student achieves an average grade less than 50%. Standard errors are clustered by school and are reported in parentheses. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

Table 15: Descriptive Statistics by University Field of Studies 2003-2011

Field of studies	**Mean Enrolment**								**Difference**	
	Girls				**Boys**					
	Mean (1)	(sd) (2)	**Mean** (3)	(sd) (4)	**Mean** (5)	(sd) (6)	**Mean** (7)	(sd) (8)	(se) (1)-(5)	(se) (3)-(7)
Exact Science	0.099	(0.298)	0.121	(0.326)	0.223	(0.416)	0.273	(0.445)	-0.125	0.152
									(0.003)	(0.003)
Science	0.046	(0.209)	0.056	(0.230)	0.037	(0.188)	0.045	(0.207)	0.009	0.011
									(0.002)	(0.002)
Social Science	0.227	(0.419)	0.278	(0.448)	0.213	(0.409)	0.260	(0.438)	0.014	0.018
									(0.003)	(0.004)
Humanities	0.273	(0.445)	0.334	(0.472)	0.088	(0.284)	0.108	(0.310)	0.184	0.226
									(0.003)	(0.004)
Vocational-non academic studies	0.172	(0.377)	0.211	(0.408)	0.258	(0.437)	0.314	(0.464)	-0.086	-0.104
									(0.003)	(0.004)
Not enrolled in post-secondary studies	0.184	(0.387)			0.181	(0.385)			0.003	
									(0.003)	

Notes: The sample includes 30,740 female students and 21,496 male students. Columns (3) and (7) refer only to enrollment in university studies. Humanities include the departments of Liberal Arts, Physcology, Journalism, Philosophy, Education, Greek Language, History, Foreign Languages, Home Economics and Law. Social Science includes the departments of Economics, Statistics, Business and Management, Accounting, Political and European studies. Exact Science includes the departments of Mathematics, Engineering, Physics and Computer Science. Science includes the departments of Biology, Chemistry, Medicine, Pharmacy, Veterinary Studies and Dentistry. Vocational-non academic studies include students who enrol in technical education institutes and agricultural studies.

Table 16: Effect of 11th and 12th Grade Own Teacher's Bias Measured in Other Classes on the Choice of University Field of Study by Gender, Sample of 21 Schools

Dependent Variable: Dummy variable for the choice of University study						
	(1)	(2)	(3)	(4)	(5)	(6)
	BOYS			GIRLS		
	11th grade (2003-2005)					
	0.029	0.029	0.037	-0.036	-0.036	-0.057
	(0.015)*	(0.016)*	(0.024)	(0.016)**	(0.016)**	(0.022)**
Sample Size	5,228	5,228	5,228	6,590	6,590	6,590
	12th grade (2003-2005)					
	-0.014	-0.015	-0.018	-0.035	-0.035	-0.034
	(0.015)	(0.015)	(0.020)	(0.015)**	(0.015)**	(0.019)*
Sample Size	4,307	4,307	4,307	5,316	5,316	5,316
	12th grade (2003-2011)					
	-0.009	-0.007	-0.013	-0.023	-0.022	-0.017
	(0.011)	(0.011)	(0.015)	(0.011)**	(0.011)**	(0.016)
Sample Size	6,650	6,650	6,650	8,270	8,270	8,270
Major FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
School FE		✓			✓	
Class FE			✓			✓

Notes: The datasets include stacked observations for each field's related subject/exam. Each row presents estimates from separate OLS regressions. All specifications include the students' blind score in 11th (Panel A) or 12th grade (Panel B and Panel C) as a control. Standard errors are clustered by class and are reported in parentheses. The dependent variable is the choice to study in Social Science, Science, Exact Science or Humanities departments. The subjects that we use for each field of study are the following: for exact science departments we use the blind score and the bias in algebra, geometry and physics in 11th grade, and mathematics and physics in 12th grade. For humanity departments we use the blind score and the bias in history and modern greek in both 11th and 12th grades. For social science departments we use the blind score and the bias in history and modern greek in 11th, and economics in 12th grade. For science departments we use the blind score and the bias in algebra, geometry and physics in 11th grade, and biology in 12th grade. We control for the 11th grade blind score in the related subjects in panel A, and for the 12th grade blind score in panels B and C. The scores are standardised and have a zero mean and a standard deviation of one. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

Table 17: Effect of 11th and 12th Grade Teacher Bias on Students' Probability to Study a Major that is the Natural Follow-up of their School Track

Dependent Variable: Dummy for enrolling in a university field that is equivalent to students' high school track				
11th grade sample 2003-2005 and 12th grade sample 2003-2011				
	BOYS	GIRLS	BOYS	GIRLS
	(1)	(2)	(3)	(4)
	11th grade		12th grade	
For those who are in the Classics Track				
(Enrol in Humanities)				
Bias in related core subjects	0.050	-0.142	0.056	-0.036
	(0.100)	(0.056)**	(0.065)	(0.032)
Sample Size	434	1,714	544	2,094
For those who are in the Science Track				
(Enrol in Science)				
Bias in related core subjects	0.127	-0.145	0.088	-0.063
	(0.073)*	(0.089)*	(0.113)	(0.089)
Sample Size	1,284	1,347	527	652
For those who are in the Exact Science Track				
(Enrol in Exact Science)				
Bias in related core subjects	-0.067	-0.132	0.010	-0.001
	(0.103)	(0.076)*	(0.047)	(0.049)
Sample Size	91	121	1,625	963
	Different Grouping of University Departments			
For those who are in the Classics Track (Enrol in Humanities and Social Science)				
Bias in related core subjects	0.029	-0.171	0.113	-0.045
	(0.104)	(0.057)***	(0.056)*	(0.036)
Sample Size	498	1,809	590	2,230
For those who are in the Science Track (Enrol in Exact Science and Science)				
Bias in related core subjects	0.123	0.126	0.140	-0.129
	(0.092)	(0.158)	(0.146)	(0.091)
Sample Size	1,484	1,449	829	1,093
For those who are in the Exact Science Track (Enrol in Exact Science and Science)				
Bias in related core subjects	-0.089	-0.147	0.017	-0.019
	(0.096)	(0.079)*	(0.039)	(0.022)
Sample Size	2,244	1,326	3,288	2,035
Year FE	✓	✓	✓	✓
School FE	✓	✓	✓	✓

Notes: Each estimate presents OLS estimates from a different regression. The outcome variable is a dummy variable that takes the value of one if students enrol in a university field that is equivalent to their high school track. All specifications include students' blind score as a control. Standard errors are clustered by school and are reported in parentheses. For the regressions run in columns (1) and (2), we include controls for the track chosen in the 12th grade. The 11th grade related biases in the core subjects are: geometry, algebra and physics for students in exact science and science tracks, and modern greek and history for students in the classics track. In 12th grade the following related biases are used: modern greek and history for students in the classics track, physics and mathematics for students in exact science track and physics, mathematics and biology for students in the science track. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

Table 18: Effect of Teacher Bias on the Probability of Enrollment in a Specific Major Over Another Conditional on the High School Track

Dependent Variable: Study a specific major over another for high school students in the same track						
	(1)	(2)	(3)	(4)	(5)	(6)
		BOYS			GIRLS	
	OLS	OLS	LOGIT	OLS	OLS	LOGIT
A: Dependent Variable: Dummy variable for studying Engineering Vs other study programs for high school students in the school science track.						
Bias in science track in 11th grade	0.048	0.006	0.013	-0.048	-0.006	-0.014
	(0.032)	(0.031)	(0.049)	(0.030)	(0.032)	(0.049)
Sample Size	3,853	3,853	3,853	4,759	4,759	4,759
Bias in science track in 12th grade	0.007	0.005	0.011	-0.001	-0.001	-0.003
	(0.003)***	(0.003)*	(0.004)**	(0.018)	(0.022)	(0.055)
Sample Size	4,254	4,254	4,254	5,790	5,790	5,790
B: Dependent Variable: Dummy variable for studying Mathematics Vs other study programs for high school students in the science track.						
Bias in science track in 11th grade	0.010	0.014	0.011	-0.001	0.002	-0.001
	(0.005)**	(0.005)***	(0.004)**	(0.003)	(0.004)	(0.004)
Sample Size	3,209	3,209	3,209	3,543	3,543	3,543
Bias in science track in 12th grade	-0.000	-0.002	0.001	0.001	0.001	-0.001
	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Sample Size	3,772	3,772	3,772	5,147	5,147	5,147
C: Dependent Variable: Dummy variable for studying Computer Science Vs other study programs for high school students in the exact science track.						
Bias in science/exact science subjects in 11th grade	0.011	0.012	0.016	-0.005	-0.006	-0.014
	(0.004)**	(0.005)**	(0.006)**	(0.004)	(0.004)	(0.011)
Sample Size	10,657	10,657	10,657	7,071	7,071	7,071
Bias in science/exact science subjects in 12th grade	0.001	0.003	0.029	-0.007	-0.009	-0.011
	(0.004)	(0.004)	(0.032)	(0.004)*	(0.004)**	(0.004)**
Sample Size	18,891	18,891	18,891	11,293	11,293	11,293
D: Dependent Variable: Dummy variable for studying Engineering or Mathematics or Computer Science Vs other study programs for high school students in the science or exact science tracks.						
Bias in science/exact science track in 11th grade	0.012	0.013	0.012	0.005	0.006	0.004
	(0.005)**	(0.006)**	(0.005)**	(0.004)	(0.005)	(0.004)
Sample Size	11,895	11,895	11,895	12,277	12,277	12,277
Bias in science/exact science track in 12th grade	-0.004	-0.003	-0.002	-0.003	-0.003	-0.002
	(0.002)*	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)
Sample Size	30,488	30,488	30,488	36,374	36,374	36,374

Notes: All specifications include a constant and the blind performance of a student in the related subjects. Standard errors are clustered by school and are reported in parentheses. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

Table 19: Effect of 11th and 12th Grade Teacher Bias on the Average Quality of the Program Students Enrolled

Dependent Variable: Percentile rank of university program								
11th grade sample 2003-2005 and 12th grade sample 2003-2011								
	Rank based on cutoffs		Rank based on mean perf.		Rank based on cutoffs		Rank based on mean perf.	
	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	11th grade				12th grade			
For those who enroll in Humanities								
Bias in related core subjects	0.212	-11.941	0.127	-9.874	7.769	-3.108	5.864	-2.920
	(6.168)	(2.168)***	(5.051)	(1.802)***	(5.190)	(1.412)**	(3.991)	(1.139)**
Sample Size	195	838	195	838	266	1,082	266	1,082
For those who enroll in Social Science								
Bias in related core subjects	-1.641	-3.561	-1.325	-2.964	3.810	-5.385	3.833	-4.027
	(4.255)	(2.758)	(3.492)	(2.267)	(3.134)	(2.840)*	(2.527)	(2.044)*
Sample Size	339	356	242	356	401	633	401	633
For those who enroll in Science								
Bias in related core subjects	27.762	-1.163	22.403	-0.953	25.325	-3.292	19.289	-1.884
	(11.348)**	(8.903)	(9.362)**	(7.389)	(12.283)*	(9.328)	(11.715)	(7.933)
Sample Size	91	121	91	121	289	396	289	396
For those who enroll in Exact Science								
Bias in related core subjects	4.261	-0.355	3.576	-0.118	0.517	3.135	0.150	2.550
	(3.733)	(3.509)	(3.076)	(2.924)	(3.905)	(4.504)	(3.207)	(3.561)
Sample Size	557	339	557	339	698	455	698	455
	11th grade				12th grade			
For those who enroll in Humanities or Social Science								
Bias in related core subjects	-0.008	-6.720	-0.023	-8.126	5.724	-4.395	5.022	-3.580
	(2.247)	(1.425)***	(2.750)	(1.729)***	(2.485)**	(2.003)**	(1.948)**	(1.641)**
Sample Size	437	1,194	437	667	1,715	667	1,715	667
For those who enroll in Science or Exact Science								
Bias in related core subjects	6.647	0.531	5.481	0.573	3.747	1.800	2.699	1.808
	(3.801)*	(3.981)	(3.142)*	(3.317)	(2.918)	(3.915)	(2.440)	(3.119)
Sample Size	648	460	648	460	987	851	987	851
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
School FE	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Each number is the OLS estimate from a different regression. Students are assigned the rank of their enrolled university department as a measure of the quality of the program they enrol. All specifications include students' blind score as a control. The rank of the university program a student enrols is a ranking based on 1) the mean admission cutoff for each university department and 2) the mean performance of enrolled students for each university department over a period of 9 cohorts (2003-2011). Standard errors are clustered by school are reported in parentheses. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

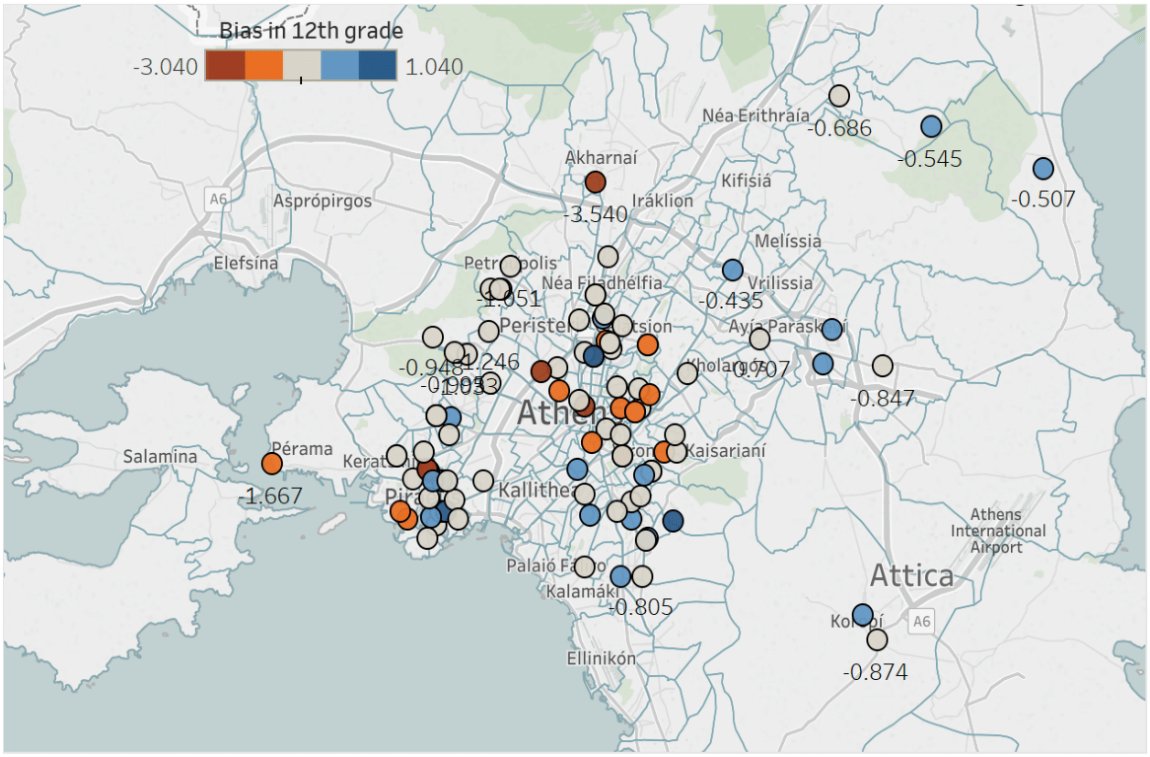
Table 20: Effect of 11th and 12th Grade Teacher Bias on the Quality of the Program Students Enrolled Based on 2003

Dependent Variable: Percentile rank of university program								
11th grade sample 2004-2005 and 12th grade sample 2004-2011								
	Rank based on mean perf. 2003		Rank based on cutoffs 2003		Rank based on mean perf. 2003		Rank based on cutoffs 2003	
	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Sample of 21 schools				
	11th grade				12th grade			
Other classes bias core	0.004	-0.049	0.347	-5.185	0.001	-0.007	1.196	-1.640
	(0.052)	(0.021)*	(3.559)	(2.664)*	(0.022)	(0.019)	(1.857)	(1.607)
Sample Size	4,940	6,049	6,274	7,488	6,161	8,265	8,016	10,492
Other classes bias classics	0.157	-0.094	11.735	-5.293	0.020	-0.005	0.418	-2.359
	(0.060)**	(0.072)	(6.051)*	(5.573)	(0.038)	(0.020)	(3.402)	(1.973)
Sample Size	360	1,352	472	1,764	5,286	8,346	6,980	10,780
Other classes bias science	-0.013	-0.064	-2.310	-8.006	-0.001	0.011	0.856	-1.813
	(0.102)	(0.074)	(9.666)	(4.245)*	(0.035)	(0.022)	(3.049)	(2.146)
Sample Size	912	1,000	996	1,076	4,782	5,821	5,909	6,889
Other classes bias in ES	0.179	-0.070	11.207	-3.879	-0.012	0.050	0.916	1.434
	(0.113)	(4.230)**	(6.743)	(0.212)	(0.022)	(0.045)	(1.909)	(4.109)
Sample Size	1,332	812	1,832	1,036	6,411	6,417	8,363	7,992
				Sample of 135 schools				
	11th grade				12th grade			
Bias in core subjects	0.035	-0.020	3.410	-2.545	0.024	-0.002	1.740	-0.274
	(0.012)***	(0.011)*	(0.907)***	(1.095)**	(0.004)***	(0.005)	(0.363)***	(0.402)
Sample Size	39,389	49,539	48,391	59,948	83,505	104,705	106,600	134,365
Bias in classics	-0.009	-0.036	3.453	-3.599	0.017	-0.009	1.771	-1.148
	(0.040)	(0.021)*	(2.901)	(1.671)**	(0.008)**	(0.006)	(0.800)**	(0.480)**
Sample Size	3,256	12,116	4,220	15,344	51,500	80,360	66,285	105,310
Bias in science	0.025	-0.007	1.955	-1.271	0.011	-0.009	2.062	-0.812
	(0.023)	(0.027)	(1.947)	(2.029)	(0.006)*	(0.007)	(0.561)***	(0.644)
Sample Size	7,620	8,892	8,068	9,396	53,870	59,790	65,520	71,600
Bias in exact science	0.047	0.015	3.222	-0.090	0.002	-0.014	1.113	-1.758
	(0.019)**	(0.025)	(1.447)**	(2.176)	(0.008)	(0.010)	(0.543)**	(0.869)**
Sample Size	11,412	6,408	15,040	8,252	65,600	55,890	84,200	69,955
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
School FE	✓	✓	✓	✓	✓	✓	✓	✓

Notes: All specifications include students' blind score as a control. The rank of the university program a student enrols is a ranking based on 1) the 2003 mean performance of enrolled students for each university department and 2) the 2003 mean performance of enrolled students for each university department. Standard errors are clustered by school are reported in parentheses. Students are assigned the rank of their enrolled university department as a measure of the quality of the program they enrol. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

Figure 1: Map of schools in the sample

Map of schools and bias in 12th grade for each school (Athens)



Map of schools and bias in 11th grade for each school (Athens)

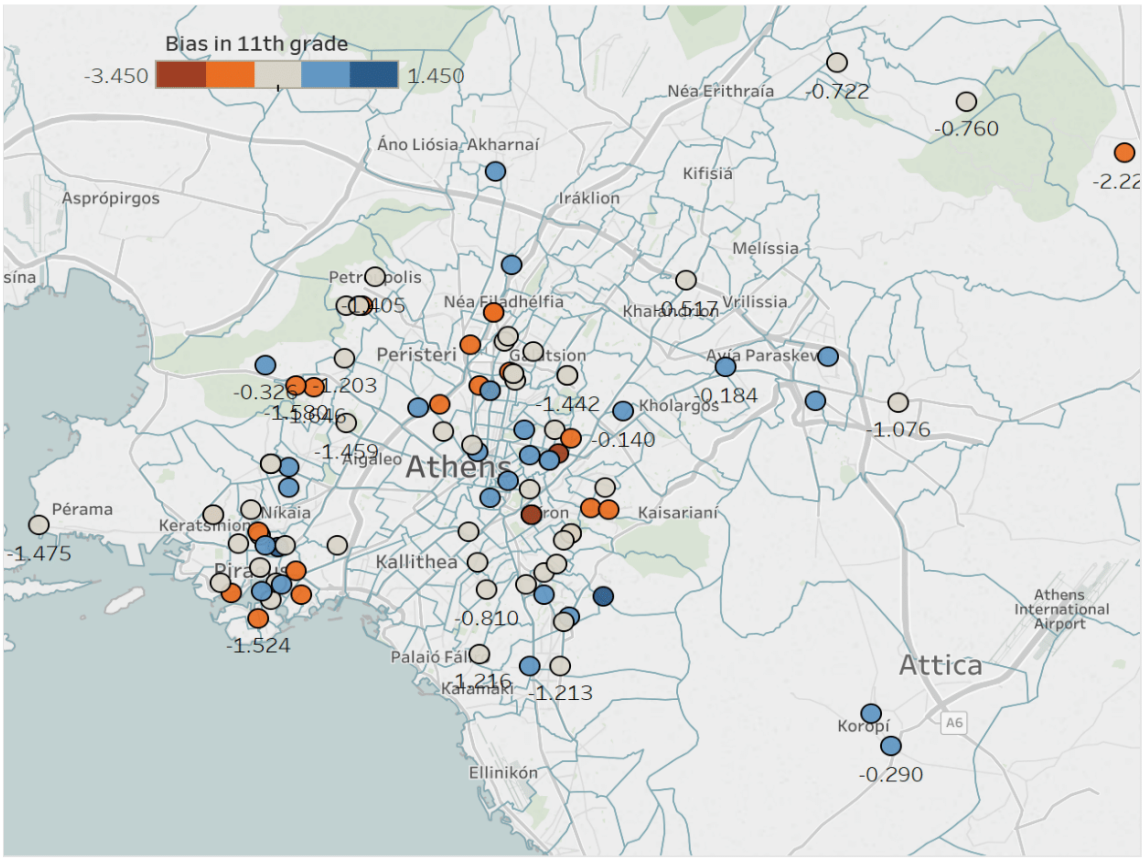


Figure 3: Distribution of Teacher Bias in Own Class and All Other Classes in 11th and 12th Grade, Sample of 21 Schools

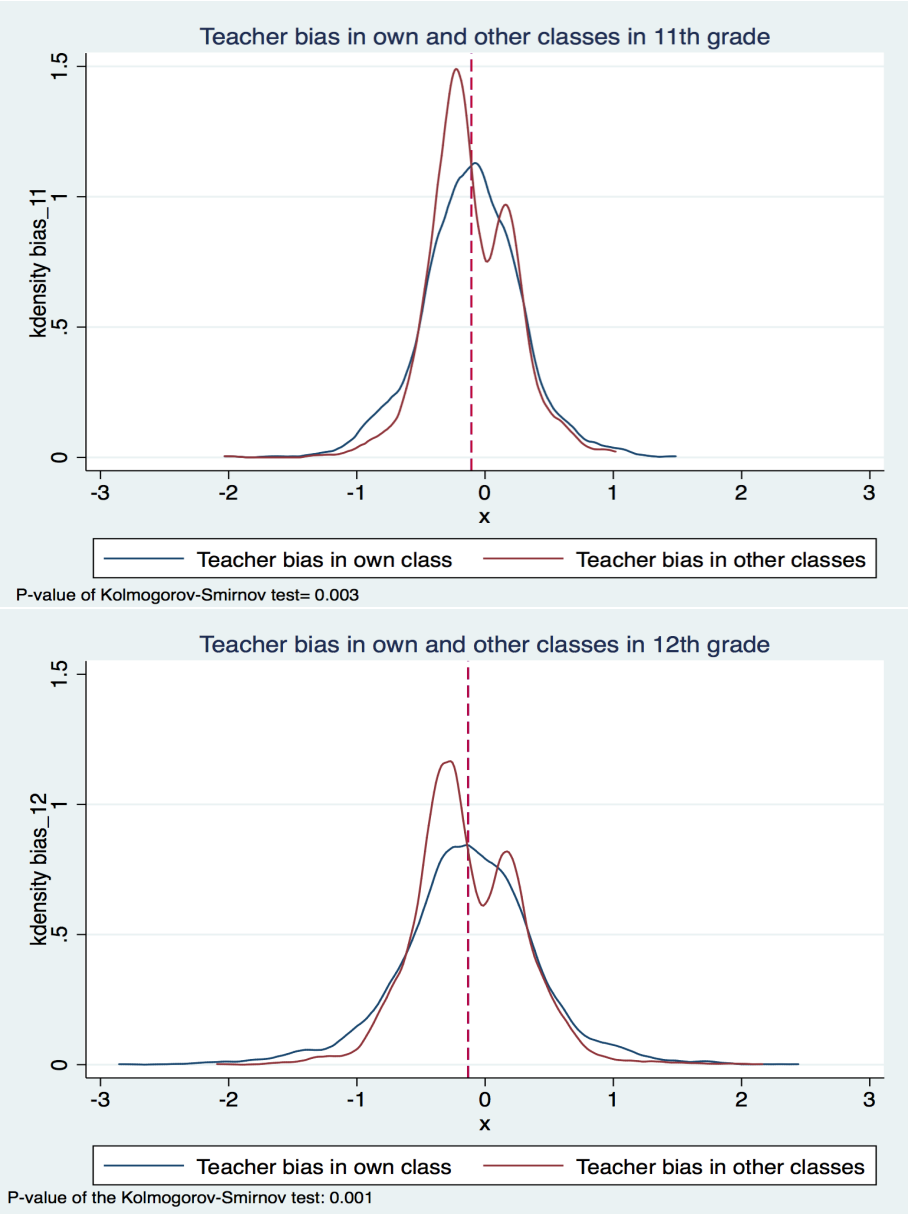


Figure 4: Distribution of Teacher Bias in All Other Classes in 11th and 12th Grade, Sample of 21 Schools

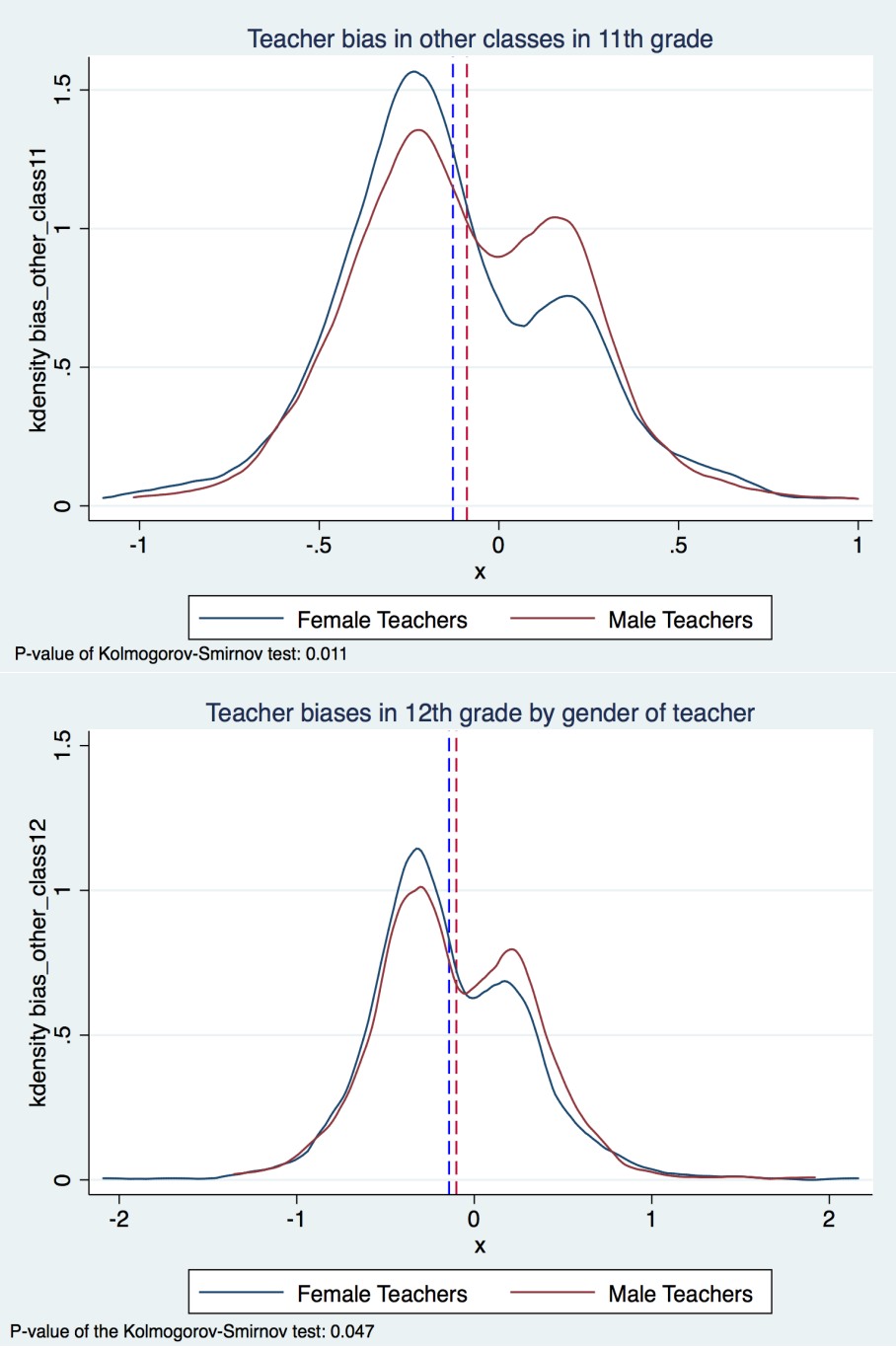


Figure 5: Average Teacher Bias in All Other Classes for Each School Track in 11th and 12th Grades, Sample of 21 Schools

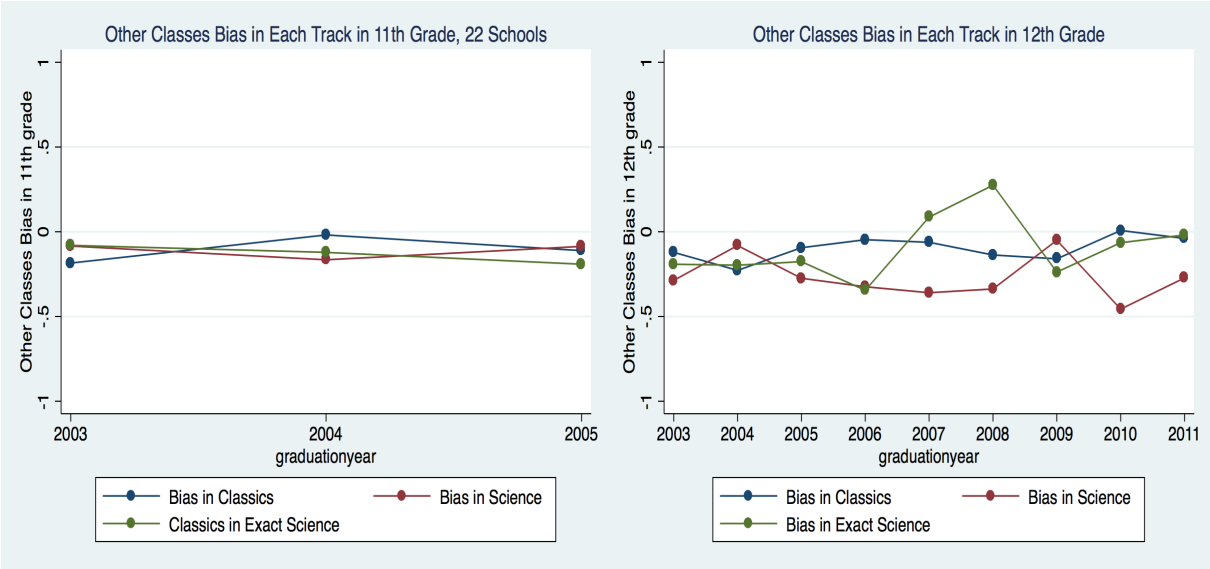


Figure 6: Average Teacher Bias by Field of Study in 11th and 12th Grades, Sample of 135 schools

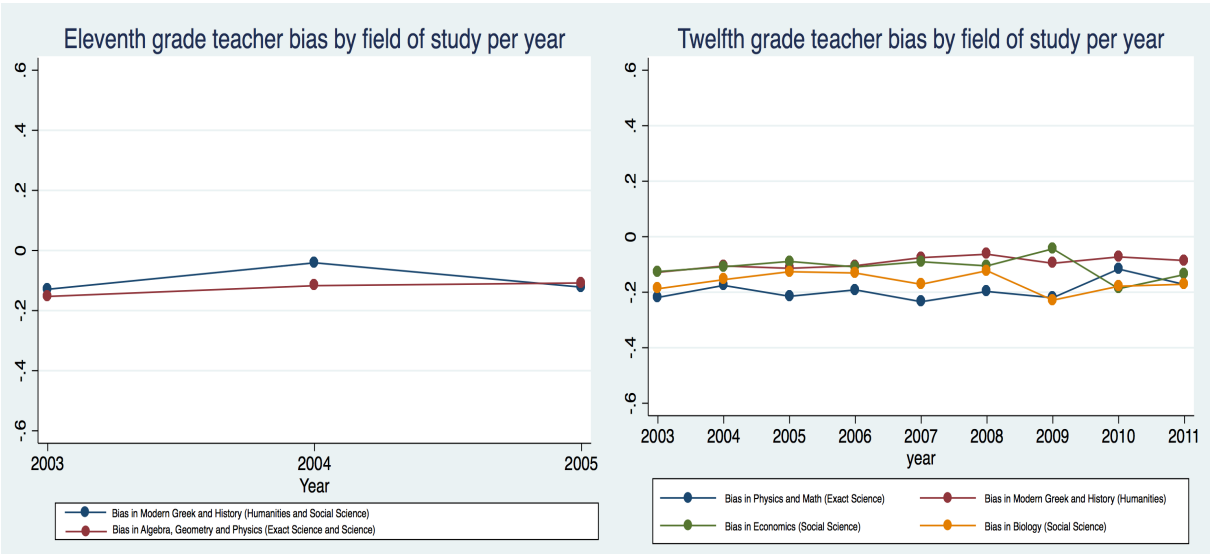


Figure 7: Proportion of Students Enrolled in Each Field at the University Level, Sample of 135 schools

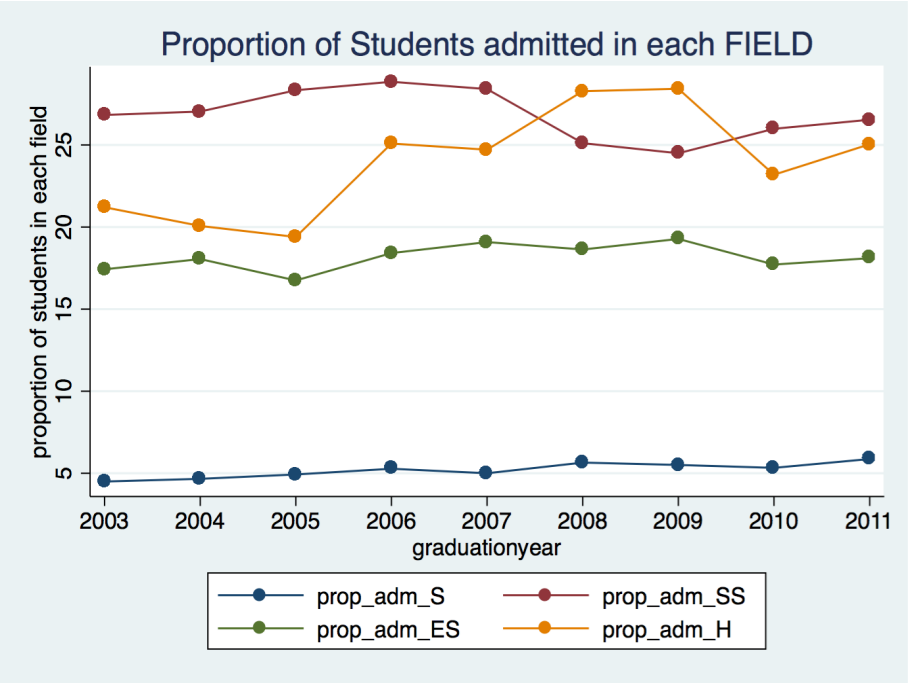


Figure 8: Proportion of Boys and Girls Enrolled in each Field at the University Level, Sample of 135 schools

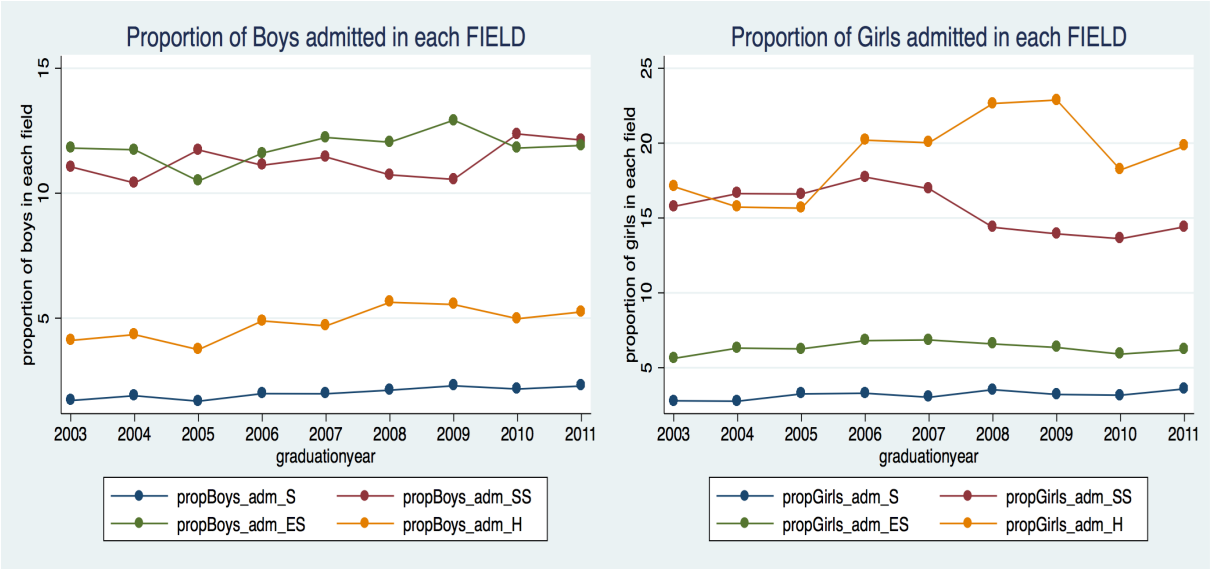
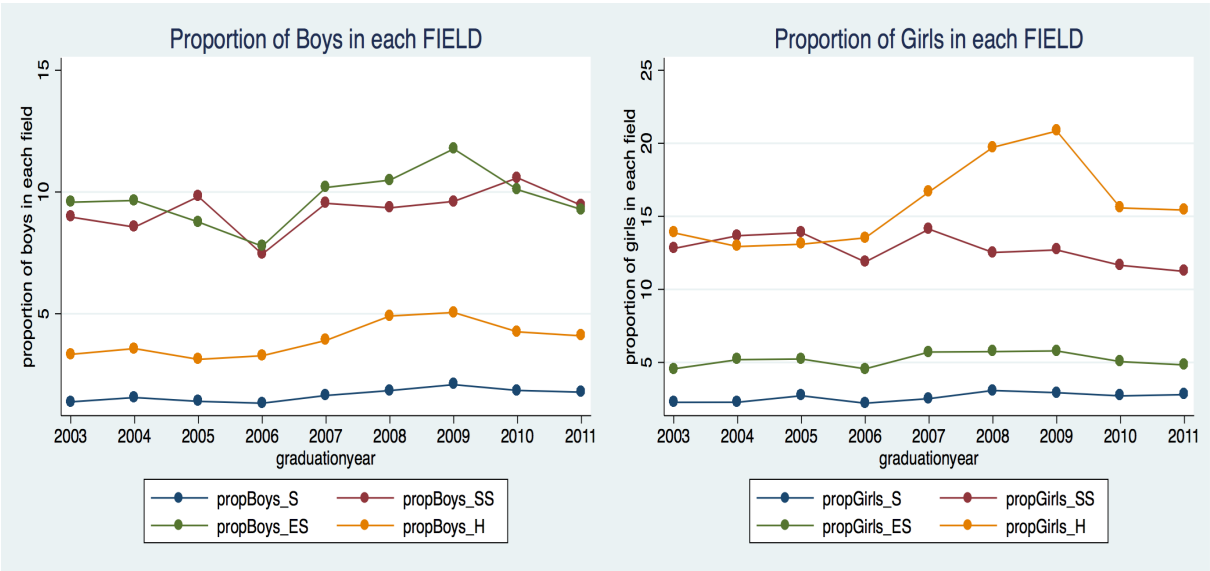


Figure 9: Proportion of Boys and Girls in Each Field at the University Level, Sample of 135 schools



Appendices

Table A1: Descriptive Statistics for Different Measures of Teacher Bias in 11th and 12th Grades for the Samples of 116 and 21 Schools

11th grade sample 2003-2005 and 12th grade sample 2003-2011						
	Teacher Bias measured in the own class (121 schools)		Teacher Bias measured in the own class (21 schools)			
Variable	**Mean**	**(sd)**	**Mean**	**(sd)**	**Diff.**	**(se)**
Bias in	(1)	(2)	(3)	(4)	(5)	(6)
11th grade						
Core subjects						
Modern Greek	-0.110	(0.415)	-0.097	(0.463)	-0.013	(0.008)
History	-0.083	(0.376)	-0.128	(0.378)	0.045	(0.007)
Algebra	-0.127	(0.303)	-0.106	(0.306)	-0.021	(0.006)
Geometry	-0.133	(0.304)	-0.094	(0.303)	-0.038	(0.006)
Physics	-0.169	(0.332)	-0.094	(0.319)	-0.075	(0.006)
Classics Track						
Ancient Greek	-0.142	(0.380)	-0.152	(0.396)	0.010	(0.012)
Philosophy	-0.172	(0.438)	-0.027	(0.415)	-0.145	(0.014)
Latin	-0.172	(0.396)	-0.087	(0.372)	-0.086	(0.013)
Science Track						
Mathematics	-0.105	(0.302)	-0.066	(0.326)	-0.039	(0.011)
Physics	-0.107	(0.369)	-0.018	(0.332)	-0.090	(0.013)
Chemistry	-0.089	(0.359)	-0.077	(0.351)	-0.012	(0.013)
Exact Science Track						
Mathematics	-0.138	(0.328)	-0.080	(0.307)	-0.058	(0.010)
Physics	-0.157	(0.324)	-0.105	(0.334)	-0.052	(0.010)
Technology and Computers	-0.316	(0.442)	-0.248	(0.397)	-0.068	(0.014)
12th grade						
Core subjects						
Modern Greek	-0.073	(0.448)	-0.134	(0.122)	0.061	(0.002)
Biology	-0.167	(0.667)	-0.141	(0.136)	-0.027	(0.003)
History	-0.160	(0.417)	-0.140	(0.117)	-0.020	(0.003)
Mathematics	-0.173	(0.539)	-0.143	(0.110)	-0.030	(0.002)
Physics	-0.247	(0.433)	-0.148	(0.101)	-0.099	(0.003)
Classics Track						
Ancient Greek	-0.084	(0.374)	-0.136	(0.111)	0.052	(0.001)
Latin	-0.084	(0.403)	-0.140	(0.096)	0.056	(0.002)
Literature	-0.115	(0.519)	-0.140	(0.112)	0.026	(0.002)
History	-0.149	(0.411)	-0.145	(0.093)	-0.004	(0.002)
Science Track						
Biology	-0.037	(0.609)	-0.143	(0.115)	0.106	(0.002)
Mathematics	-0.177	(0.503)	-0.149	(0.119)	-0.028	(0.002)
Physics	-0.221	(0.573)	-0.152	(0.091)	-0.069	(0.002)
Chemistry	-0.118	(0.596)	-0.146	(0.126)	0.028	(0.002)
Exact Science Track						
Mathematics	-0.145	(0.299)	-0.142	(0.090)	-0.003	(0.001)
Physics	-0.224	(0.331)	-0.149	(0.090)	-0.075	(0.001)
Business Administration	-0.154	(0.384)	-0.143	(0.100)	-0.011	(0.001)
Computers	-0.198	(0.326)	-0.148	(0.087)	-0.050	(0.001)
Optional						
Economics	-0.121	(0.388)	-0.140	(0.099)	0.020	(0.001)

Notes: Negative bias means that the teacher is pro-girl. The means are weighted by number of students.

Table A2: Gender Differences in Teacher Bias

Dependent Variable: Bias measured in other classes in any year						
	11th grade			12th grade		
	(1)	(2)	(3)	(4)	(5)	(6)
Core Subjects						
Constant	-0.089	-0.069	-0.052	-0.129	-0.098	-0.098
	(0.020)***	(0.041)*	(0.038)	(0.020)***	(0.034)***	(0.036)***
Female teacher dummy	-0.047	-0.049	-0.068	0.032	0.019	0.013
	(0.028)*	(0.026)*	(0.027)**	(0.032)	(0.036)	(0.037)
Sample Size	710	710	710	663	663	663
Classics Subjects						
Constant	-0.093	-0.055	-0.047	-0.129	-0.107	-0.103
	(0.022)***	(0.035)	(0.032)	(0.021)***	(0.032)***	(0.035)***
Female teacher dummy	-0.033	-0.040	-0.052	0.022	0.024	0.018
	(0.030)	(0.027)	(0.029)*	(0.031)	(0.035)	(0.036)
Sample Size	928	928	928	738	738	738
Science Subjects						
Constant	-0.089	-0.066	-0.052	-0.134	-0.108	-0.104
	(0.020)***	(0.041)*	(0.038)	(0.021)***	(0.032)***	(0.035)***
Female teacher dummy	-0.046	-0.049	-0.068	0.036	0.026	0.020
	(0.028)*	(0.026)*	(0.027)**	(0.031)	(0.035)	(0.036)
Sample Size	711	711	711	737	737	737
Exact Science Subjects						
Constant	-0.086	-0.059	-0.048	-0.125	-0.095	-0.110
	(0.018)***	(0.041)	(0.038)	(0.019)***	(0.030)***	(0.033)***
Female teacher dummy	-0.058	-0.061	-0.076	0.021	0.006	0.016
	(0.026)**	(0.024)**	(0.025)***	(0.028)	(0.030)	(0.031)
Sample Size	810	810	810	1,018	1,018	1,018
No controls	✓			✓		
Subjects FE		✓	✓		✓	✓
Year FE		✓	✓		✓	✓
School FE			✓			✓

Notes: The sample of 21 schools is used. All specifications include a constant. Standard errors are clustered by class and are reported in parentheses. *,**,*** denotes significance at the 10%,5% and 1% level respectively. A teacher sample is used here.

Table A3: Descriptive Statistics for Teachers in 11th and 12th Grade, Sample of 21 schools

11th grade sample 2003-2005 and 12th grade sample 2003-2011						
	Bias measured in own class					
	Male Teachers		Female Teachers			
	Mean	(s.d.)	Mean	(s.d.)	Difference	(se)
	11th grade					
Core Subjects						
Modern Greek	-0.074	(0.425)	-0.143	(0.426)	0.069	(0.099)
Algebra	-0.057	(0.425)	-0.150	(0.257)	0.092	(0.098)
Geometry	-0.066	(0.372)	-0.090	(0.300)	0.025	(0.088)
Physics	0.052	(0.325)	0.002	(0.354)	0.051	(0.086)
History	-0.179	(0.482)	-0.095	(0.451)	-0.085	(0.109)
Track:Classics						
Ancient Greek	-0.075	(0.282)	-0.191	(0.521)	0.116	(0.151)
Philosophy	-0.083	(0.554)	-0.145	(0.537)	0.062	(0.179)
History	-0.154	(0.400)	-0.164	(0.892)	0.010	(0.170)
Track: Science						
Mathematics	-0.008	(0.494)	-0.212	(0.418)	0.204	(0.162)
Physics	0.063	(0.348)	-0.193	(0.502)	0.256	(0.151)
Chemistry	0.053	(0.614)	-0.043	(0.357)	0.096	(0.212)
Track: Exact Science						
Mathematics	0.025	(0.325)	-0.177	(0.370)	0.202	(0.121)
Physics	-0.045	(0.382)	-0.128	(0.357)	0.083	(0.126)
Technology	-0.213	(0.470)	-0.120	(0.724)	-0.093	(0.217)
	12th grade					
Core Subjects						
Modern Greek	0.035	(0.599)	-0.239	(0.533)	0.274	(0.109)
History	-0.074	(0.463)	-0.181	(0.535)	0.107	(0.121)
Mathematics	-0.156	(0.630)	-0.112	(0.498)	-0.043	(0.143)
Physics	-0.081	(0.411)	-0.178	(0.350)	0.097	(0.106)
Biology	-0.150	(0.595)	0.029	(0.625)	-0.180	(0.141)
Track:Classics						
Ancient Greek	-0.153	(0.589)	-0.299	(0.564)	0.147	(0.157)
Latin	0.052	(0.609)	-0.173	(0.413)	0.224	(0.131)
Literature	-0.095	(0.562)	-0.134	(0.639)	0.039	(0.155)
History	-0.201	(0.514)	-0.236	(0.427)	0.036	(0.123)
Track: Science						
Biology	-0.316	(1.142)	0.134	(0.477)	-0.450	(0.396)
Mathematics	-0.192	(0.732)	0.008	(0.873)	-0.200	(0.297)
Physics	-0.269	(0.602)	-0.209	(0.715)	-0.061	(0.227)
Chemistry	-0.239	(0.972)	-0.154	(0.621)	-0.393	(0.332)
Track: Exact Science						
Mathematics	-0.097	(0.472)	-0.056	(0.391)	-0.041	(0.133)
Physics	-0.215	(0.462)	-0.129	(0.235)	-0.086	(0.122)
Bus. Administ.	-0.171	(0.506)	0.108	(0.714)	-0.279	(0.162)
Computers	-0.288	(0.532)	-0.209	(0.879)	-0.080	(0.200)
Optional						
Economics	0.001	(0.365)	-0.190	(0.680)	0.192	(0.156)

Notes: A teacher sample is used here. This table presents the means, the standard deviations and the differences of teacher bias for male and female teachers in 11th and 12th grades.

Table A4: Mean Scores and Standard Deviations in the National Exam (blind) and the School Exam (non-blind) in 11th Grade, 2003-2005, Sample of 143 Schools

	NATIONAL			SCHOOL		
Variables	Boys	Girls	**Difference**	Boys	Girls	**Difference**
	(sd)	(sd)	(se)	(sd)	(sd)	(se)
	(1)	(2)	(1)-(2)	(3)	(4)	(3)-(4)
Core subjects						
Modern Greek	-0.204	0.165	-0.369	-0.258	0.209	-0.467
	(0.997)	(0.972)	(0.013)	(1.016)	(0.936)	(0.013)
History	-0.116	0.094	-0.210	-0.162	0.131	-0.293
	(0.964)	(0.997)	(0.018)	(1.016)	(0.967)	(0.013)
Physics	0.058	-0.047	0.104	-0.046	-0.037	-0.083
	(1.002)	(0.995)	(0.013)	(1.024)	(0.978)	(0.013)
Algebra	0.037	-0.030	0.068	-0.047	0.038	-0.085
	(1.010)	(0.990)	(0.013)	(1.023)	(0.980)	(0.013)
Geometry	0.046	-0.037	0.084	-0.040	0.032	-0.072
	(1.000)	(0.999)	(0.013)	(1.017)	(0.985)	(0.013)
Track: Classics						
Ancient Greek	-0.234	0.062	-0.296	-0.321	0.085	-0.406
	(1.003)	(0.983)	(0.027)	(1.065)	(0.964)	(0.027)
Philosophy	-0.181	0.048	-0.229	-0.284	0.075	-0.360
	(0.985)	(0.998)	(0.027)	(1.083)	(0.963)	(0.027)
Latin	-0.241	0.064	-0.305	-0.335	0.089	-0.424
	(1.012)	(0.987)	(0.027)	(1.089)	(0.955)	(0.027)
Track: Science						
Mathematics	0.010	-0.009	0.019	-0.050	0.046	-0.096
	(1.014)	(0.987)	(0.024)	(1.003)	(0.968)	(0.024)
Physics	0.028	-0.026	0.054	-0.018	0.016	-0.034
	(0.996)	(1.002)	(0.024)	(1.005)	(0.995)	(0.024)
Chemistry	-0.003	0.003	-0.006	-0.046	0.042	-0.089
	(1.000)	(1.000)	(0.024)	(1.007)	(0.991)	(0.024)
Track: Exact Science						
Mathematics	-0.051	0.094	-0.146	-0.100	0.184	-0.284
	(0.993)	(1.006)	(0.022)	(1.005)	(0.963)	(0.022)
Physics	-0.034	0.064	-0.098	-0.082	0.152	-0.235
	(0.997)	(1.001)	(0.022)	(1.011)	(0.960)	(0.022)
Computer Science	0.025	-0.046	0.071	-0.085	0.157	-0.241
	(0.972)	(1.049)	(0.022)	(1.004)	(0.973)	(0.022)

Notes: The national and school exams scores are standardized z-scores. This table presents test scores gender gaps by type of exam (blind and non-blind) and subject in 12th grade. A positive difference means that boys outperform girls in that type of exams, while a negative difference means that girls outperform boys. The sample includes 23,608 students. The non-blind score in each subject is the score in the second term school exam.

Table A5: Mean Scores and Standard Deviations in the National Exam (blind) and the School Exam (non-blind) in 12th Grade, 2003-2011, Sample of 135 Schools

	National Exam			School Exam		
Variables	Boys (sd)	Girls (sd)	**Difference** (se)	Boys (sd)	Girls (sd)	**Difference** (se)
	(1)	(2)	(1)-(2)	(3)	(4)	(3)-(4)
Modern Greek	-0.204 (1.000)	0.158 (0.970)	-0.363 (0.008)	-0.240 (1.056)	0.186 (0.912)	-0.426 (0.008)
History	-0.065 (0.984)	0.052 (1.010)	-0.117 (0.013)	-0.147 (1.034)	0.119 (0.955)	-0.266 (0.013)
Mathematics	0.034 (1.002)	-0.034 (0.997)	0.068 (0.009)	-0.052 (1.052)	0.052 (0.942)	-0.104 (0.009)
Physics	0.085 (1.002)	-0.070 (0.993)	0.155 (0.013)	-0.048 (1.037)	-0.039 (0.967)	-0.087 (0.013)
Biology	-0.040 (0.992)	0.024 (1.004)	-0.064 (0.010)	-0.143 (1.099)	0.086 (0.925)	-0.229 (0.010)
Track: Classics						
Ancient Greek	-0.224 (1.035)	0.060 (0.981)	-0.284 (0.015)	-0.277 (1.117)	0.075 (0.952)	-0.352 (0.015)
Latin	-0.256 (1.047)	0.066 (0.976)	-0.308 (0.015)	-0.301 (1.117)	0.081 (0.950)	-0.382 (0.015)
Modern Literature	-0.239 (1.045)	0.065 (0.977)	-0.304 (0.015)	-0.326 (1.148)	0.088 (0.937)	-0.414 (0.015)
History	-0.063 (0.988)	0.017 (1.002)	-0.112 (0.016)	-0.180 (1.098)	0.049 (0.966)	-0.229 (0.016)
Track: Science						
Biology	-0.003 (1.013)	0.002 (0.990)	-0.005 (0.020)	-0.029 (1.041)	0.020 (0.969)	-0.049 (0.020)

Continued on next page

Table A5 – *Continued from previous page*

	NATIONAL			SCHOOL		
Variables	Boys (sd) (1)	Girls (sd) (2)	**Difference** (se) (1)-(2)	Boys (sd) (3)	Girls (sd) (4)	**Difference** (se) (3)-(4)
Mathematics	0.123	-0.085	0.208	0.006	-0.004	0.009
	(1.008)	(0.985)	(0.020)	(1.041)	(0.970)	(0.020)
Physics	0.159	-0.111	0.270	-0.009	-0.106	0.025
	(0.988)	(0.993)	(0.020)	(1.011)	(0.991)	(0.020)
Chemistry	0.080	-0.055	0.135	0.002	0.002	-0.004
	(0.979)	(1.010)	(0.020)	(1.022)	(0.984)	(0.020)
Track: Exact Science						
Mathematics	-0.026	0.044	-0.070	-0.087	0.146	-0.233
	(1.016)	(0.970)	(0.012)	(1.037)	(0.916)	(0.012)
Physics	0.012	-0.021	0.033	-0.076	0.128	-0.204
	(1.021)	(0.964)	(0.012)	(1.035)	(0.924)	(0.012)
Business Administration	-0.066	0.110	-0.176	-0.131	0.219	-0.351
	(0.996)	(0.997)	(0.012)	(1.058)	(0.850)	(0.012)
Application Development	0.006	-0.010	0.015	-0.074	0.123	-0.196
	(1.012)	(0.980)	(0.012)	(1.048)	(0.900)	(0.012)
Optional						
Economics	-0.024	0.023	-0.047	-0.087	0.083	-0.171
	(0.995)	(1.004)	(0.011)	(1.046)	(0.947)	(0.011)

Note: The national and school exams scores are standardized z-scores. This table presents test scores gender gaps by type of exam (blind and non-blind) and subject in 12th grade. A positive difference means that boys outperform girls in that type of exams while a negative difference means that girls outperform boys. The sample includes 64,650 students. The non-blind score in each subject is the score in the second term school exam.

Table A6: Effect of 11th Grade Teacher Bias Measured in the Own Class on the Blind Score in 12th Grade, Sample of 21 Schools

Dependent Variable: Blind score in 12th grade national exams						
	BOYS			GIRLS		
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Sample of 21 Schools					
Core Subjects						
Own teacher's 11th grade bias	0.107	0.123	0.157	-0.150	-0.124	-0.146
	(0.034)***	(0.030)***	(0.032)***	(0.035)***	(0.033)***	(0.035)***
Sample Size	9,745	9,745	9,745	12,264	12,264	12,264
Classics Subjects						
Own teacher's 11th grade bias	0.075	0.099	0.099	-0.124	-0.098	-0.108
	(0.074)	(0.055)*	(0.058)*	(0.040)***	(0.035)***	(0.036)***
Sample Size	1,892	1,892	1,892	7,305	7,305	7,305
Science Subjects						
Own teacher's 11th grade bias	0.134	0.142	0.170	-0.115	-0.114	-0.109
	(0.054)**	(0.049)***	(0.053)***	(0.044)**	(0.041)***	(0.043)**
Sample Size	3,334	3,334	3,334	3,485	3,485	3,485
Exact Science Subjects						
Own teacher's 11th grade bias	0.065	0.065	0.119	-0.149	-0.134	-0.182
	(0.039)*	(0.037)*	(0.040)***	(0.062)**	(0.059)**	(0.060)***
Sample Size	5,869	5,869	5,869	3,460	3,460	3,460
	Panel B: Sample of 135 Schools					
Core Subjects						
Own teacher's 11th grade bias	0.130	0.130	0.142	-0.099	-0.099	-0.126
	(0.012)***	(0.011)***	(0.012)***	(0.012)***	(0.010)***	(0.011)***
Sample Size	75,315	75,315	75,315	92,798	92,798	92,798
Classics Subjects						
Own teacher's 11th grade bias	0.101	0.121	0.139	-0.098	-0.097	-0.115
	(0.021)***	(0.019)***	(0.020)***	(0.013)***	(0.011)***	(0.013)***
Sample Size	18,973	18,973	18,973	70,331	70,331	70,331
Science Subjects						
Own teacher's 11th grade bias	0.121	0.118	0.135	-0.064	-0.077	-0.108
	(0.017)***	(0.015)***	(0.016)***	(0.015)***	(0.013)***	(0.014)***
Sample Size	29,793	29,793	29,793	34,375	34,375	34,375
Exact Science Subjects						
Own teacher's 11th grade bias	0.104	0.100	0.099	-0.118	-0.120	-0.120
	(0.012)***	(0.011)***	(0.011)***	(0.017)***	(0.015)***	(0.015)***
Sample Size	62,395	62,395	62,395	33,664	33,664	33,664
Subjects FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
School FE		✓	✓		✓	✓
Class FE			✓			✓

Notes: The datasets for the core subjects and each track include stacked observations for each subject/exam. Each row presents estimates from separate OLS regressions. All specifications include the students' blind score as a control. The second panel "Classics Subjects" includes relevant exams from the core and the classics track. The third panel "Science Subjects" includes relevant exams from the core and the science track. The forth panel "Exact Science Subjects" includes relevant exams from the core and the exact science track. Standard errors are clustered by class and are reported in parentheses. All scores are standardised z-scores. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

Table A7: Number and Proportion of Students Enrolled in Different University Studies by School Tracks

			All	Notenrolled	Human.	Science	Exact Science	Social Science	Vocational
			(1)	(2)	(3)	(4)	(5)	(6)	(7)
Track11	Classics		8,323	1,862	3,179	41	232	1,643	1,364
		%		22.3	38.2	0.5	2.8	19.7	16.4
	Science		6,821	371	390	788	1,990	1,538	1,752
		%		5.4	5.7	11.6	29.2	22.5	25.7
	Exact Science		8,843	1,986	425	103	1,329	2,236	2,766
		%		22.5	4.8	1.2	15.0	25.3	31.3
2003-2011									
Track12	Classics		24,622	5,808	10,936	84	530	4,302	2,959
		%		23.6	44.4	0.3	2.2	17.5	12.0
	Science		10,587	563	483	2,239	2,929	1,292	3,094
		%		5.3	4.6	21.1	27.7	12.2	29.2
	Exact Science		30,712	5,639	1,324	474	6,650	8,881	5,639
		%		18.4	4.3	1.5	21.7	28.9	18.4
2003-2005									
Track12	Classics		8,367	1,878	3,211	39	224	1,653	1,360
		%		22.4	38.4	0.5	2.7	19.8	16.3
	Science		4,265	213	223	744	1,293	666	1,134
		%		5.0	5.2	17.4	30.3	15.6	26.6
	Exact Science		11,433	2,146	571	149	2,053	3,118	3,399
		%		18.8	5.0	1.3	18.0	27.3	29.7

Note: Column (2) presents the number and the proportion of students who are not enrolled in any university department. Column (3) ,(4), (5), (6), (7) presents the number and the proportion of students who enroll in humanities, science, exact science, social science and vocational studies respectively. Note: Cohorts 2003 2004 2005 used.

		Pathway12		
		Classics	Science	Exact Science
Pathway11	Classics	10,233	15	68
	Science	43	4,844	2,841
	Exact Science	76	53	10,201

Note: Cohorts 2003 2004 2005 used.

Table A8: Effect of 11th Grade Teacher Bias on Enrollment in University Schooling

Dependent Variable: Dummy variable for the decision of student to enrol in University						
	(1)	(2)	(3)	(4)	(5)	(6)
	BOYS			GIRLS		
	OLS	OLS	LOGIT	OLS	OLS	LOGIT
Bias in all subjects in 11th grade	0.004	0.005	0.006	-0.007	-0.007	-0.003
	(0.002)*	(0.002)**	(0.001)***	(0.002)***	(0.002)***	(0.002)
Sample Size	10,405	10,405	10,405	12,561	12,561	12,561
Bias in core subjects in 11th grade	0.004	0.006	0.007	-0.010	-0.009	-0.005
	(0.003)	(0.003)*	(0.002)***	(0.003)**	(0.003)**	(0.002)**
Sample Size	10,604	10,604	10,604	13,054	13,054	13,054
Bias in Classics subjects in 11th grade	0.019	0.021	0.027	-0.004	-0.000	-0.001
	(0.013)	(0.016)	(0.011)**	(0.006)	(0.008)	(0.005)
Sample Size	1,696	1,696	1,696	6,179	6,179	6,179
Bias in Science subjects in 11th grade	0.001	0.004	0.002	-0.005	-0.014	-0.001
	(0.007)	(0.007)	(0.004)	(0.005)	(0.007)*	(0.004)
Sample Size	3,219	3,219	3,219	3,549	3,549	3,549
Bias in Exact Science subjects in 11th grade	0.009	0.008	0.010	-0.042	-0.041	-0.032
	(0.002)	(0.003)	(0.006)	(0.011)***	(0.014)***	(0.009)***
Sample Size	5,674	5,674	5,674	3,071	3,071	3,071
Bias in all subjects in 12th grade	0.007	0.006	0.007	-0.007	-0.008	-0.004
	(0.002)***	(0.002)**	(0.001)***	(0.001)***	(0.001)***	(0.001)***
Sample Size	28,323	28,323	28,323	35,269	35,269	35,269
Bias in core subjects in 12th grade	0.001	0.000	0.008	-0.003	-0.003	-0.003
	(0.003)	(0.003)	(0.002)***	(0.002)	(0.003)	(0.002)*
Sample Size	10,812	10,812	10,812	13,284	13,284	13,284
Bias in classics subjects in 12th grade	0.013	0.017	0.010	-0.003	-0.000	-0.001
	(0.004)**	(0.006)**	(0.003)**	(0.002)	(0.002)	(0.002)
Sample Size	1,696	1,696	1,696	6,179	6,179	6,179
Bias in science subjects in 12th grade	0.009	0.009	0.002	-0.008	-0.008	-0.002
	(0.02)***	(0.004)**	(0.001)*	(0.002)***	(0.002)***	(0.001)
Sample Size	4,254	4,254	4,254	5,790	5,790	5,790
Bias in exact science subjects in 12th grade	0.015	0.013	0.010	-0.021	-0.022	-0.006
	(0.003)***	(0.003)***	(0.002)***	(0.004)***	(0.004)***	(0.002)**
Sample Size	18,891	18,891	18,891	11,293	11,293	11,293
Year FE	✓	✓	✓	✓	✓	✓
School FE		✓	✓		✓	✓

Notes: All specifications include a constant and the blind performance of a student in the corresponding subjects. Columns (3) and (6) repost the marginal effects of the logistic regressions. Standard errors are clustered by class are are reported in parentheses. The outcome variable is a dummy variable that takes the value of one if the student enrols in university and zero otherwise. The variable "Blind score in all subjects in 11th grade" is the average blind score in all subjects in 11th grade. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

Table A9: Effect of 11th and 12th Grade Own Teacher's Bias Measured in the Own Class on the Choice of University Field of Study by Gender, Sample of 21 Schools

Dependent Variable: Dummy variable for the choice of University study						
	(1)	(2)	(3)	(4)	(5)	(6)
	BOYS			GIRLS		
	11th grade (2003-2005)					
	0.013	0.014	0.018	-0.044	-0.046	-0.054
	(0.016)	(0.016)	(0.024)	(0.012)***	(0.012)***	(0.017)***
Sample Size	5,588	5,588	5,588	7,036	7,036	7,036
	12th grade (2003-2005)					
	-0.011	-0.011	-0.007	-0.029	-0.024	-0.020
	(0.013)	(0.015)	(0.021)	(0.013)**	(0.014)*	(0.017)
Sample Size	4,307	4,307	4,307	5,316	5,316	5,316
	12th grade (2003-2011)					
	0.007	0.008	0.002	-0.017	-0.015	-0.018
	(0.009)	(0.010)	(0.014)	(0.009)*	(0.009)*	(0.013)
Sample Size	6,650	6,650	6,650	8,270	8,270	8,270
Major FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
School FE		✓			✓	
Class FE			✓			✓

Notes: The datasets include stacked observations for each field's related subject/exam. Each row presents estimates from separate OLS regressions. All specifications include the students' blind score in 11th (Panel A) or 12th grade (Panel B and Panel C) as a control. Standard errors are clustered by class and are reported in parentheses. The dependent variable includes the choice to study in Social Science, Science, Exact Science and Humanities departments. The subjects that we use for each field of study are the following: for exact science departments we use the blind score and the bias in algebra, geometry and physics in 11th grade, and mathematics and physics in 12th grade. For humanity departments we use the blind score and the bias in history and modern greek in both 11th and 12th grades. For social science departments we use the blind score and the bias in history and modern greek in 11th, and economics in 12th grade. For science departments we use the blind score and the bias in algebra, geometry and physics in 11th grade, and biology in 12th grade. We control for the 11th grade blind score in the related subjects in panel A, and for the 12th grade blind score in panels B and C. The scores are standardised and have a zero mean and a standard deviation of one. The scores are standardised and have a zero mean and a standard deviation of one. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

Table A10: The Effect of the Proportion of Female Teachers in the Class of Core Subjects, in the Track and in the School on Teachers Bias

Dependent Variable: Bias of own teacher in all other classes				
	11th grade		**12th grade**	
Core Subjects				
Proportion of female teacher in the (core subjects) class	-0.008	-0.120	-0.034	-0.047
	(0.072)	(0.089)	(0.043)	(0.048)
Sample Size	1,130	1,130	1,421	1,421
Proportion of female teacher in the school	-0.001	-0.081	-0.070	-0.099
	(0.088)	(0.136)	(0.053)	(0.071)
Sample Size	1,335	1,335	1,995	1,995
Only Classics Track Subjects				
Proportion of female teacher in the track	-0.193	-0.276	-0.128	-0.162
	(0.090)**	(0.090)***	(0.091)	(0.089)*
Sample Size	231	231	376	376
Only Science Track Subjects				
Proportion of female teacher in the track	-0.254	-0.253	-0.172	-0.110
	(0.133)*	(0.136)*	(0.217)	(0.183)
Sample Size	143	143	228	228
Only Exact Science Track Subjects				
Proportion of female teacher in the track	-0.139	-0.155	-0.121	-0.193
	(0.098)	(0.097)	(0.129)	(0.131)
Sample Size	131	131	437	437
Subjects FE	✓	✓	✓	✓
Year FE		✓		✓

Notes: Each row presents estimates from separate OLS regressions. Standard errors are clustered by class and are reported in parentheses. *,**,*** denotes significance at the 10%,5% and 1% level respectively. A teacher sample is used here.

Table A11: Effect of 11th and 12th Grade Teacher Bias on the Quality of the Program Students Enrolled, Sample of 135 Schools

Dependent Variable: Percentile rank of university program a student enrols				
	Rank based on degree cutoffs in terms of the national exam scores of enrolled students		Rank based on mean exam score of enrolled students national exam score of enrolled students	
	BOYS	GIRLS	BOYS	GIRLS
	(1)	(2)	(3)	(4)
			11th grade	
Bias in core subjects	0.713	-0.385	0.662	-0.439
	(0.155)***	(0.189)**	(0.129)***	(0.143)***
Sample Size	8,749	10,765	10,604	13,055
Bias in classics subjects	0.853	-0.242	1.992	0.130
	(0.992)	(0.555)	(0.764)**	(0.409)
Sample Size	1,299	4,843	1,719	6,213
Bias in science subjects	0.924	-0.312	0.826	-0.378
	(0.427)**	(0.498)	(0.385)**	(0.390)
Sample Size	3,054	3,347	3,219	3,347
Bias in exact science subjects	0.386	-1.751	0.726	-1.383
	(0.491)	(0.491)***	(0.385)*	(0.469)***
Sample Size	4,367	2,404	5,657	3,071
			12th grade	
Bias in core subjects	0.276	-0.014	0.066	0.107
	(0.142)*	(0.134)	(0.142)	(0.113)
Sample Size	22,993	28,171	28,048	34,388
Bias in classics subjects	0.698	-0.251	0.557	-0.044
	(0.265)***	(0.117)**	(0.198)***	(0.086)
Sample Size	3,913	14,143	5,220	18,336
Bias in science subjects	0.318	-0.128	0.085	-0.032
	(0.151)**	(0.122)	(0.096)	(0.088)
Sample Size	4,067	5,469	4,254	5,790
Bias in exact science subjects	0.065	-0.215	0.066	-0.175
	(0.186)	(0.199)	(0.093)	(0.128)
Sample Size	15,281	9,366	18,891	11,293
Year FE	✓	✓	✓	✓
School FE	✓	✓	✓	✓

Notes: Students are assigned the rank of their enrolled university department as a measure of the quality of the program they enrol. All specifications include students' blind performance related subjects as a control. The rank of the university program a student enrols is based on a) the average cut-off for each university department in terms of the national exam scores of the enrolled students and b) the mean national exam performance of enrolled students over a period of 9 cohorts (2003-2011). Standard errors are clustered by class and are reported in parentheses. *,**,*** denotes significance at the 10%, 5% and 1% level respectively.

Table A12: Effect of 11th and 12th Grade Own Teacher's Bias on the Choice of University Study, Sample of 135 Schools

Dependent Variable: Dummy variable for the choice of University study						
	(1)	(2)	(3)	(4)	(5)	(6)
	BOYS			GIRLS		
	Panel A: 11th grade, including non-admitted students 2003-2005					
Own teacher's bias	0.011	0.014	0.014	-0.018	-0.017	-0.021
	(0.005)**	(0.006)**	(0.006)**	(0.005)***	(0.005)***	(0.008)***
Sample Size	70,307	70,307	70,307	87,945	87,945	87,945
	Panel B: 11th grade, excluding non-admitted students 2003-2005					
Own teacher's bias	0.011	0.009	0.010	-0.019	-0.021	-0.026
	(0.005)**	(0.005)*	(0.005)*	(0.005)***	(0.005)***	(0.008)***
Sample Size	41,996	41,996	41,996	51,606	51,606	51,606
	Panel C: 12th grade, including non-admitted students 2003-2011					
Own teacher's bias	0.011	0.012	0.011	-0.013	-0.013	-0.008
	(0.003)***	(0.003)***	(0.004)***	(0.003)***	(0.003)***	(0.003)**
Sample Size	109,172	109,172	109,172	133,576	133,576	133,576
	Panel D: 12th grade, excluding non-admitted students 2003-2011					
Own teacher's bias	0.010	0.009	0.007	-0.013	-0.014	-0.019
	(0.003)***	(0.003)***	(0.004)*	(0.003)***	(0.003)***	(0.004)***
Sample Size	80,861	80,861	80,861	97,237	97,237	97,237
	Panel E: 12th grade, including non-admitted students 2003-2005					
Own teacher's bias	0.010	0.014	0.016	-0.011	-0.010	-0.012
	(0.005)**	(0.005)***	(0.007)**	(0.005)**	(0.005)*	(0.006)*
Sample Size	48,798	48,798	48,798	59,710	59,710	59,710
	Panel G: 12th grade, excluding non-admitted students 2003-2005					
Own teacher's bias	0.010	0.010	0.014	-0.011	-0.010	-0.016
	(0.005)**	(0.005)**	(0.008)*	(0.005)**	(0.005)**	(0.007)**
Sample Size	38,210	38,210	38,210	46,668	46,668	46,668
Major FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
School FE		✓			✓	
Class FE			✓			✓

Note: The datasets include stacked observations for each subject/exam. All specifications include students' blind score as a control. Standard errors are clustered by class and are reported in parentheses. The dependent variable is the choice to study in Social Science, Science, Exact Science or Humanities departments. The subjects that we use for each field of study are the following: for exact science departments we use the blind score and the bias in algebra, geometry and physics in 11th grade, and mathematics and physics in 12th grade. For humanity departments we use the blind score and the bias in history and modern greek in both 11th and 12th grades. For social science departments we use the blind score and the bias in history and modern greek in 11th, and economics in 12th grade. For science departments we use the blind score and the bias in algebra, geometry and physics in 11th grade, and biology in 12th grade. We control for the 11th grade blind score in the related subjects in panels A and B and for the 12th grade blind score in panels C, D, E and G. The scores are standardised and have a zero mean and a standard deviation of one. *,**,*** denotes significance at the 10%,5% and 1% level respectively.

Figure A.1: Distribution of of Teachers Bias in Own and other Classes (in the Same Year) in 11th and 12th grade, Sample of 21 schools

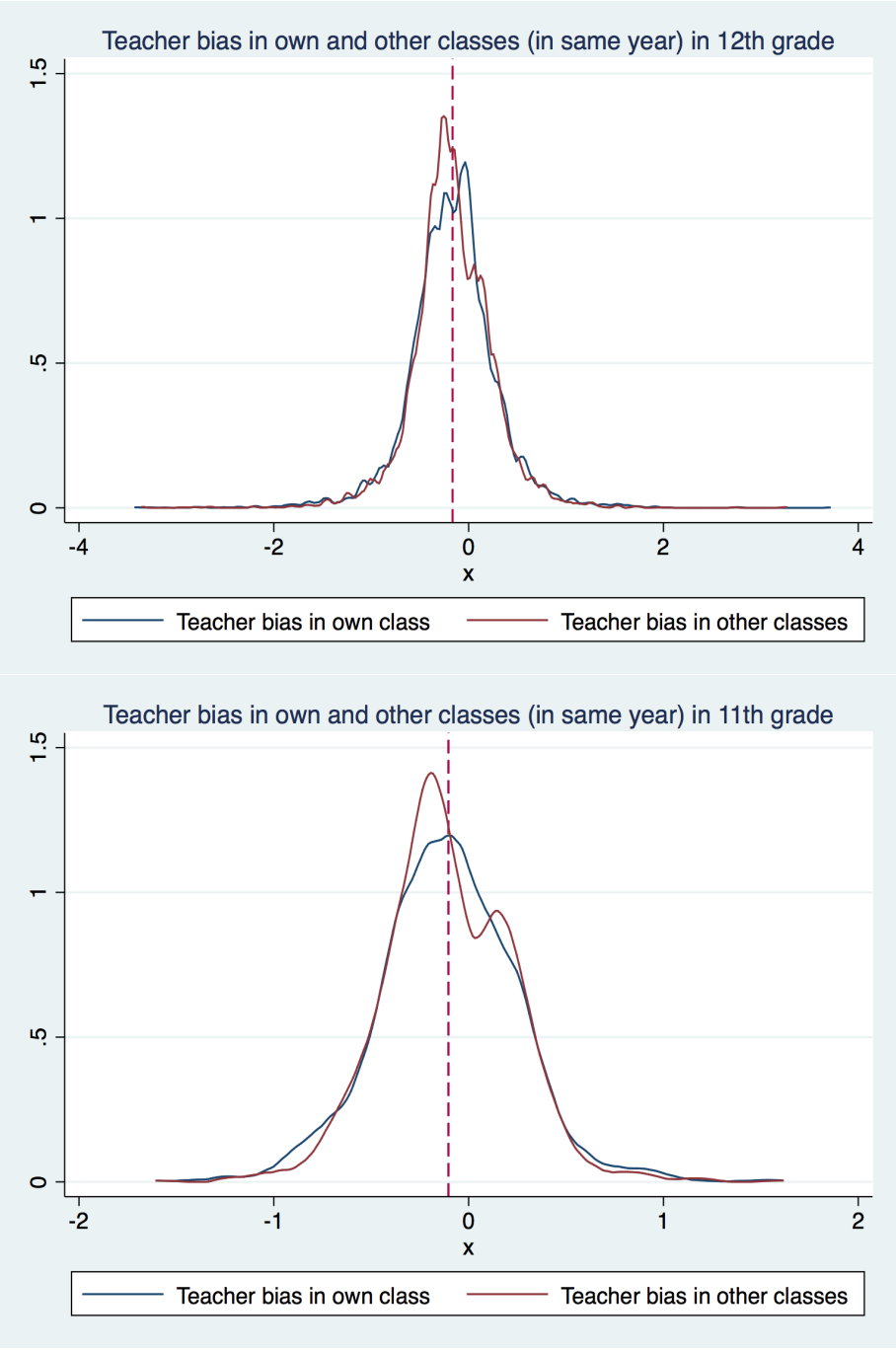


Figure A.2: Distribution of Teachers Bias in Own Class for Male and Female Teachers in 11th and 12th Grade, Sample of 21 Schools

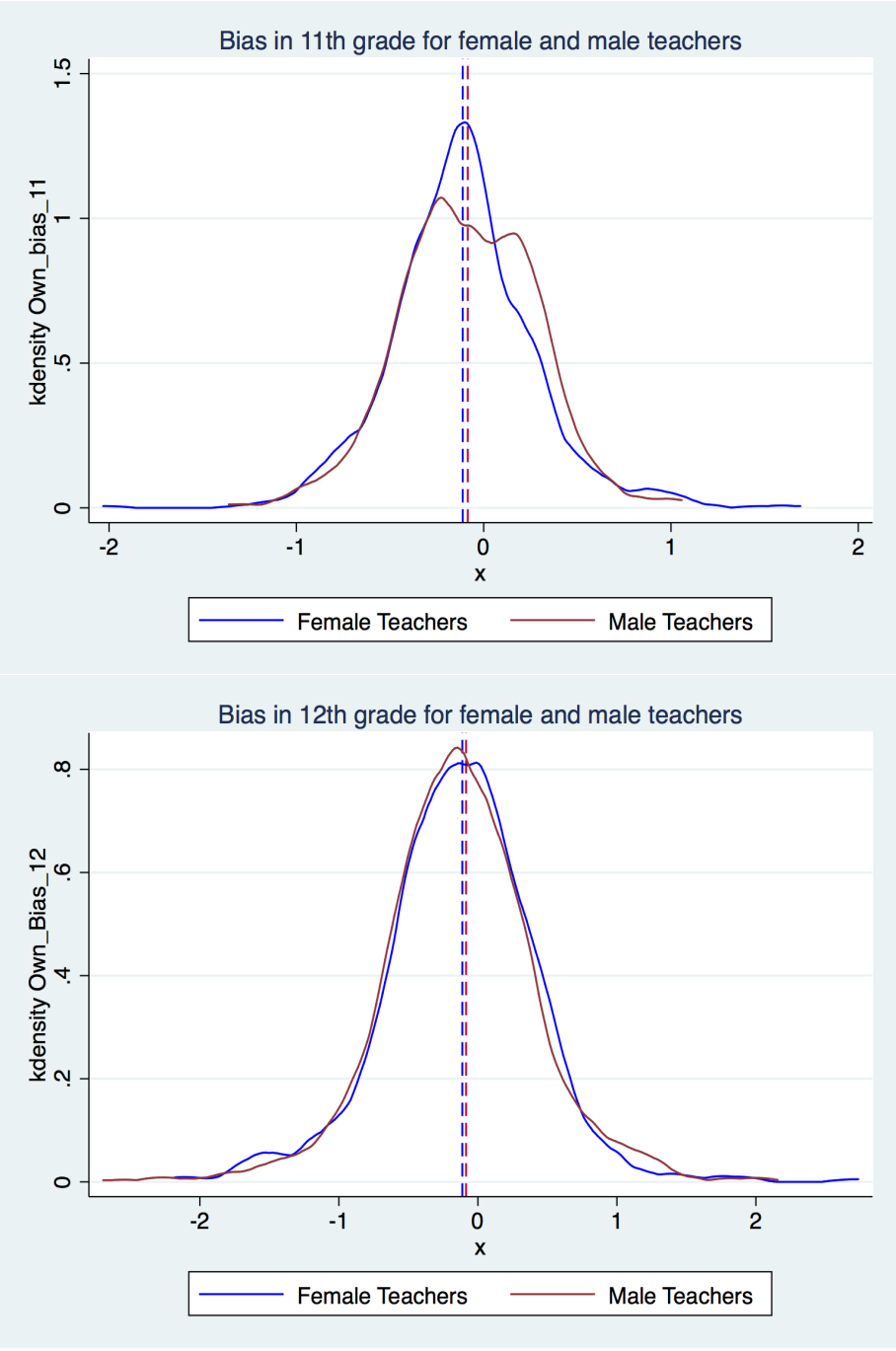


Figure A.3: Average Teacher Bias in Own Class for Each Track in 11th and 12th Grades, Sample of 135 Schools

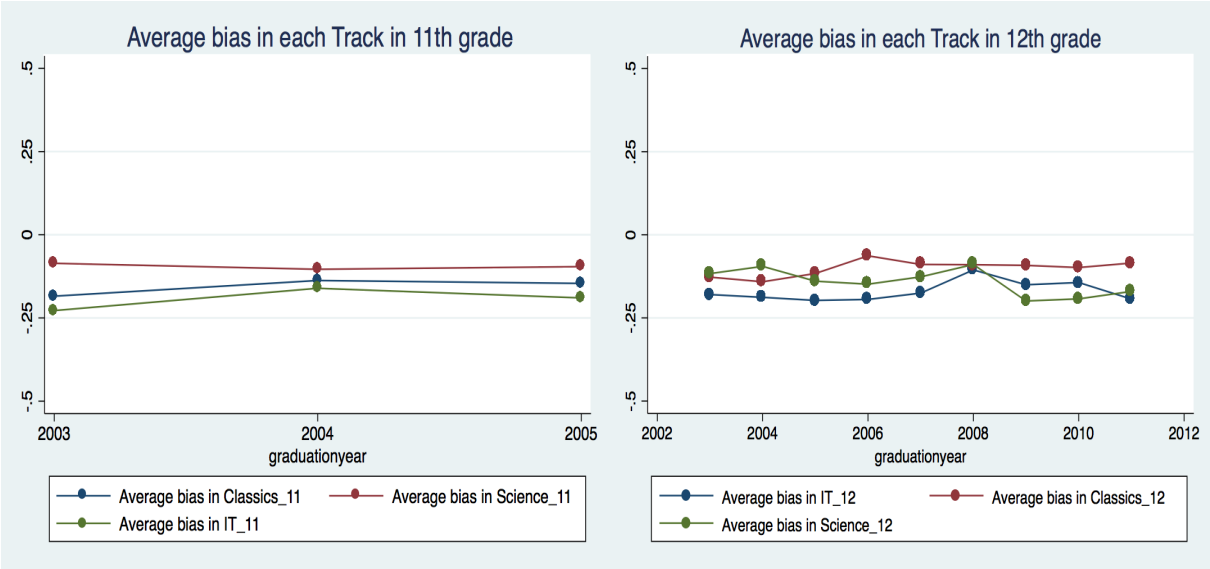


Figure A.4: Average Teacher Bias in Own Class for Each School Track in 11th and 12th Grades, Sample of 21 Schools