Online Appendix

Comparative Advantage and Gender Gap in STEM

Sofoklis Goulas	Silvia Griselda	Rigissa Megalokonomou
Hoover Institution	AXA Research Lab	Department of Economics
Stanford University	Bocconi University	University of Queensland

Descriptive Evidence of the Effect of STEM and Non-STEM Performance

Panel A of Table A1 focuses on the subgroup of students who enroll in a non-STEM (columns (1) and (6)) and STEM track in 11^{th} grade (columns (2) and (7)), separately. As shown in columns (1) and (2), males who enroll in a STEM track in 11^{th} grade have a higher STEM performance but slightly lower non-STEM performance in 10^{th} grade than males who go into non-STEM in 11^{th} grade. The corresponding differences and *p*-values are shown in columns (3) and (4), respectively. Females who enroll in a STEM track in grade 11 (column (7)) have a higher performance in both types of subjects compared with females who go into a non-STEM track (column (6)). The corresponding differences in column (8) are both positive and statistically significant (column (9)). Females going into STEM and non-STEM tracks have a higher performance in both types of subjects than males going into STEM and non-STEM tracks, respectively. The classroom average grade in STEM and non-STEM is very similar for males and females. Males and females who enroll into a STEM track have a higher absolute STEM advantage.

Panel B of Table A1 focuses on the subgroup of students who apply for a non-STEM (columns (1) and (6)) and STEM university degree (columns (2) and (7)), separately. Comparing columns (1) to (6) and (2) to (7), we find that females who apply to both types of degree programs outperform males in both types of subjects. Also, males and females who apply to a STEM degree program have a higher 10^{th} grade performance in both types of subjects compared with those who do not apply to a non-STEM degree program. Differences in classroom average grades in the two types of subjects between the two groups of university applicants are small (0.305 and 0.193 for males in STEM and non-STEM, respectively, and 0.118 and 0.086 for females in STEM and non-STEM, respectively). Males and females who apply for a STEM university degree have a higher absolute STEM advantage.

Table A1: **Descriptive Statistics by Gender and Enrollment**

	Male				Female			
Panel A	Non-STEM Track Enrollment in Grade 11	STEM Track Enrollment in Grade 11	Diff.	*p-value*	Non-STEM Track Enrollment in Grade 11	STEM Track Enrollment in Grade 11	Diff.	*p-value*
	(1)	(2)	(3)	(4)	(6)	(7)	(8)	(9)
Own Grade in STEM	8.019	11.088	3.069	0.000	9.035	12.528	3.494	0.000
Own Grade in non-STEM	12.787	12.414	-0.373	0.000	13.821	14.093	0.273	0.000
Comparative STEM Advantage	0.278	0.547	0.270	0.000	0.296	0.547	0.251	0.000
Class Av. Grade in STEM	10.084	10.286	0.202	0.000	10.115	10.319	0.204	0.000
Class Av. Grade in non-STEM	13.005	12.908	-0.098	0.000	12.918	12.892	-0.025	0.117
Own Absolute Adv. in STEM	0.622	0.891	0.269	0.000	0.645	0.884	0.240	0.000
Class Absolute Adv. in STEM	0.781	0.802	0.022	0.000	0.789	0.807	0.018	0.000
Obs	6,185	26,725			21,177	18,925		
Panel B	Non-STEM University Application	STEM University Application	Diff.	*p-value*	Non-STEM University Application	STEM University Application	Diff.	*p-value*
Own Grade in STEM	8.717	11.972	3.256	0.000	11.202	13.147	1.944	0.000
Own Grade in non-STEM	10.884	12.975	2.091	0.000	13.232	14.482	1.251	0.000
Comparative STEM Advantage	0.454	0.583	0.129	0.000	0.492	0.573	0.081	0.000
Class Av. Grade in STEM	10.063	10.368	0.305	0.000	10.241	10.359	0.118	0.000
Class Av. Grade in non-STEM	12.766	12.959	0.193	0.000	12.833	12.919	0.086	0.001
Own Absolute Adv. in STEM	0.799	0.926	0.126	0.000	0.840	0.906	0.066	0.000
Class Absolute Adv. in STEM	0.794	0.806	0.012	0.000	0.806	0.808	0.002	0.378
Obs	7,058	19,523			5,560	13,188		

Notes: This table shows summary statistics for student's own performance in STEM and non-STEM subjects in grade 10, classroom average performance in STEM and non-STEM subjects in grade 10, own and classroom absolute STEM performance in grade 10 for different subgroups by gender, separately. Panel A reports these statistics for students who enroll in a non-STEM and STEM tracks in grade 11 for males and females, separately. Panel B reports these statistics for students who apply for a non-STEM and STEM university degree for males and females, separately. Columns (3) and (8) report the differences and columns (4) and (9) report the *p*-values for the *t*-test on the difference between non-STEM and STEM enrollment.

Effect of Absolute STEM Advantage on STEM Study

In this section, we examine empirically the following hypothesis: The higher an individual's competence in STEM relative to their competence in non-STEM, the more likely they are to specialize in STEM, while controlling for his/her peers' competence in STEM relative to non-STEM. Similarly, the higher a female student's peers' competence in STEM relative to non-STEM, the less likely she is to specialize in STEM, while keeping constant her competence in STEM relative to non-STEM.

An individual's competence in STEM relative to their competence in non-STEM can be proxied using definition (4). A similar definition can be used to proxy a student's peers' competence in STEM relative to non-STEM. We investigate the association between own and peer advantage in STEM using the following specification:

$$Y_{ijt} = \beta_0 + \beta_1 \underbrace{\frac{Grade\_STEM_{ijt}}{Grade\_nonSTEM_{ijt}}}_{\text{STEM advantage}} + \beta_2 \underbrace{\frac{Av\_Classroom\_Grade\_STEM_{ijt}}{Av\_Class\_Grade\_nonSTEM_{ijt}}}_{\text{Classroom STEM advantage}} + \mu_{st} + \varepsilon_{ijt} \tag{A1}$$

Table A2 presents our estimates of model (A1). Higher (absolute) STEM advantage increases the likelihood to enroll in a STEM track in grade 11. Moreover, (absolute) STEM advantage is positively correlated with the likelihood of applying for a STEM university degree program.

Table A2: **Association between Students' Own and Classroom STEM Advantage on STEM Study Outcomes**

	STEM Track in Grade 11		Applied for STEM University Degree	
	(1)	(2)	(3)	(4)
Female	-0.287***	-0.433***	-0.024***	0.037
	(0.004)	(0.024)	(0.005)	(0.027)
Abs. STEM Advantage	0.652***	0.448***	0.274***	0.285***
	(0.008)	(0.009)	(0.011)	(0.012)
Abs. STEM Advantage × Female		0.430***		-0.035**
		(0.012)		(0.018)
Class Abs. STEM Advantage	-0.299***	-0.173***	-0.081***	-0.064**
	(0.025)	(0.029)	(0.028)	(0.031)
Class Abs. STEM Advantage × Female		-0.245***		-0.037
		(0.031)		(0.036)
Obs.	72,943	72,943	45,269	45,269
School x Year FE	Yes	Yes	Yes	Yes
Controls	No	No	No	No
Mean Y	0.63	0.63	0.72	0.72
St. Dev Y	0.48	0.48	0.45	0.45
Raw Gender Gap Y	-0.34	-0.34	-0.03	-0.03

Notes: This table examines the patterns of track choice in grade 11 and university departments application. Importantly, the table has no intent to identify causal inference, but rather questions whether the gender gap in STEM enrollment can be explained by gender difference in student performance. The table reports the results of a specification in which the track enrollment and university application decisions of student i, in school j, cohort t are regressed on their own and classmates' average absolute STEM advantage, school-by-cohort FE, and student's characteristics, such as gender and year of birth. Each regression includes school-year FE. Standard errors are clustered at the school cohort level. The last row in each panel shows the slope coefficient of the regression of each outcome variable on a female indicator, reflecting the gender gap in that outcome. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A3: **Association between Cardinal Comparative STEM Advantage and Future Study Decisions**

	STEM Track in Grade 11		Applied for STEM University Degree	
	(1)	(2)	(3)	(4)
Cardinal Comparative STEM Adv.	0.499***	0.346***	0.209***	0.217***
	(0.006)	(0.007)	(0.008)	(0.009)
Cardinal Comparative STEM Adv. × Female		0.317***		-0.025*
		(0.010)		(0.014)
Obs.	72,940	72,940	45,259	45,259
Classroom FE	Yes	Yes	Yes	Yes
Controls	No	No	No	No
Mean Y	0.63	0.63	0.72	0.72
St. Dev Y	0.48	0.48	0.45	0.45
Raw Gender Gap Y	-0.34	-0.34	-0.03	-0.03

Notes: This table explores the patterns of track choice in grade 11 and university department application. Importantly, the table has no intent to identify causal inference, but rather questions whether the gender gap in STEM specialization can be explained by gender differences in student performance. The table reports the results of a specification in which the track enrollment and university application decisions of student i, in classroom j, in school s, in cohort t are regressed on their cardinal comparative STEM advantage (as defined in the LHS of equation (1)). The regression includes classroom FE and student's characteristics, such as gender and year of birth. Standard errors are clustered at the school cohort level. The last row in each panel shows the slope coefficient of the regression of each outcome variable on a female indicator, reflecting the gender gap in that outcome. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A4: **Effect of Comparative STEM Advantage on STEM Study Using School-Year FE**

	Linear (1)	Quadratic (2)	Cubic (3)	Quartic (4)	Quintic (5)	Nonlinear (6)
	STEM Track in Grade 11					
Comparative STEM Advantage	0.163***	0.037	0.075***	0.065***	0.060**	0.066***
	(0.019)	(0.024)	(0.023)	(0.024)	(0.024)	(0.025)
Comparative STEM Advantage × Female	0.060**	0.201***	0.137***	0.112***	0.116***	0.098***
	(0.030)	(0.033)	(0.033)	(0.034)	(0.035)	(0.034)
Obs.	72,943	72,943	72,943	72,943	72,943	72,943
Mean of Y	0.63	0.63	0.63	0.63	0.63	0.63
St. Dev. Y	0.48	0.48	0.48	0.48	0.48	0.48
Raw Gender Gap Y	-0.34	-0.34	-0.34	-0.34	-0.34	-0.34
	Application for STEM University Degree					
Comparative STEM Advantage	0.072***	-0.057**	-0.062**	-0.060**	-0.052*	-0.026
	(0.025)	(0.028)	(0.029)	(0.030)	(0.030)	(0.031)
Comparative STEM Advantage × Female	0.085*	0.150***	0.150***	0.152***	0.143***	0.114**
	(0.044)	(0.046)	(0.046)	(0.048)	(0.048)	(0.049)
Obs.	45,269	45,269	45,269	45,269	45,269	45,269
Mean of Y	0.72	0.72	0.72	0.72	0.72	0.72
St. Dev. Y	0.45	0.45	0.45	0.45	0.45	0.45
Raw Gender Gap Y	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03
School × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated effects of comparative STEM advantage on track and university degree choices. The outcome variables are: (a) an indicator for whether a student enrolls in a STEM track in grade 11 (top panel), and (b) an indicator for whether a student applies for a STEM university degree two years later (bottom panel). For each of the two outcomes, we run different specifications for different degrees of polynomials for the absolute STEM advantage (columns (1)-(5)) and a nonlinear specification using binary indicators for each decile of absolute STEM advantage (column (6)). Each regression controls for student gender, absolute STEM advantage, STEM, non-STEM performance, classroom average STEM and non-STEM performance, classroom average absolute STEM advantage, class size, classroom gender mix, interactions of individual terms with gender, and school-year FE. Standard errors are clustered at the school cohort level. The last row in each panel shows the slope coefficient of the regression of each outcome variable on a female indicator, reflecting the gender gap in that outcome. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A5: **Effect of Comparative STEM Advantage on STEM Study Controlling for Dispersion in STEM Advantage**

	Linear	Quadratic	Cubic	Quartic	Quintic	Nonlinear
	(1)	(2)	(3)	(4)	(5)	(6)
	STEM Track in Grade 11					
Comparative STEM Advantage	0.162***	0.041*	0.077***	0.066***	0.061**	0.067***
	(0.019)	(0.024)	(0.023)	(0.024)	(0.024)	(0.025)
Comparative STEM Advantage × Female	0.059**	0.199***	0.140***	0.115***	0.118***	0.099***
	(0.029)	(0.033)	(0.033)	(0.034)	(0.034)	(0.034)
Obs.	72,940	72,940	72,940	72,940	72,940	72,940
Mean of Y	0.63	0.63	0.63	0.63	0.63	0.63
St. Dev. Y	0.48	0.48	0.48	0.48	0.48	0.48
Raw Gender Gap Y	-0.34	-0.34	-0.34	-0.34	-0.34	-0.34
	Application for STEM University Degree					
Comparative STEM Advantage	0.065**	-0.056**	-0.062**	-0.057*	-0.050*	-0.025
	(0.026)	(0.028)	(0.028)	(0.030)	(0.030)	(0.031)
Comparative STEM Advantage × Female	0.094**	0.147***	0.146***	0.143***	0.135***	0.108**
	(0.044)	(0.046)	(0.046)	(0.049)	(0.048)	(0.049)
Obs.	45,265	45,265	45,265	45,265	45,265	45,265
Mean of Y	0.72	0.72	0.72	0.72	0.72	0.72
St. Dev. Y	0.45	0.45	0.45	0.45	0.45	0.45
Raw Gender Gap Y	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03
School × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated effects of comparative STEM advantage on track and university degree choices. The outcome variables are: (a) an indicator for whether a student enrolls in a STEM track in grade 11 (top panel), and (b) an indicator for whether a student applies for a STEM university degree two years later (bottom panel). For each of the two outcomes, we run different specifications for different degrees of polynomials for the absolute STEM advantage (columns (1)-(5)) and a nonlinear specification using binary indicators for each decile of absolute STEM advantage (column (6)). Each regression controls for student gender, absolute STEM advantage, STEM, non-STEM performance, classroom average STEM and non-STEM performance, classroom average absolute STEM advantage, standard deviation of absolute STEM advantage in the classroom, interaction between the average and the standard deviation of absolute STEM advantage in the classroom, class size, classroom gender mix, interactions of individual terms with gender, and school-year FE. Standard errors are clustered at the school cohort level. The last row in each panel shows the slope coefficient of the regression of each outcome variable on a female indicator, reflecting the gender gap in that outcome. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Alternative Approaches to Controlling for STEM and non-STEM Performance

We have explored the association between performance rank in STEM and non-STEM subjects and our measure of comparative STEM advantage in the context of our empirical investigation. In particular, we have augmented our main specification to include the student's percentile ranks in STEM and non-STEM subjects. The augmented specification is as follows:

$$
\begin{aligned}
Y_{ijst} = {} & \alpha + \beta Comparative\ STEM\ Advantage_{ijst} + f(a_{ijst}) \\
& + Rank\ in\ STEM_{ijst} + Rank\ in\ non\text{-}STEM_{ijst} \\
& + X'_{ijst}\ \gamma + \mu_{jst} + \varepsilon_{ijst}
\end{aligned} \tag{A2}
$$

where Y_{ijst} is the outcome variable for i student, in j classroom, in s school, and t cohort. Outcome Y depends on comparative STEM advantage, a flexible function of absolute STEM advantage, $f(a_{ijst})$, the percentile within classroom rank in STEM and non-STEM subjects, Rank in STEM$_{ijst}$ and Rank in non-STEM$_{ijst}$ respectively, individual characteristics, X_{ijst}, and classroom FE, μ_{jst}. As in the main analysis, we model $f(a_{ijst})$ in many ways, but our preferred specification controls for absolute STEM advantage nonlinearly using 10 indicators for a student's decile position in the sample-wide distribution of absolute STEM advantage. In every specification with female interactions, every regressor is interacted with the female dummy.

The estimates from specification (A2) are reported in Table A6. We find that our parameter of interest, the coefficient of the comparative STEM advantage remains qualitatively unaffected to the inclusion of the ranks in STEM and non-STEM subjects. This suggests that our measure of comparative advantage picks an influence distinct of that the rank in STEM and non-STEM subjects pick up.

Consistent with the results found by Delaney and Devereux (2021b), the effect of classroom rank in STEM (non-STEM) subjects has positive (negative) and significant effect on male and female decisions to enroll into a STEM track in Grade 11. This result helps appease concerns that performance in non-STEM subjects may not matter for

students in Greece. In is important to highlight that the negative effect of classroom rank in non-STEM subjects is larger in magnitude for female students. This suggests that having a high rank value with respect to non-STEM performance has a higher impact on females' decision to enroll in a non-STEM track than it does for males.

Table A6: **Effect of Comparative STEM Advantage on STEM Study Controlling for Rank in STEM and Non-STEM Subjects**

	Linear (1)	Quadratic (2)	Cubic (3)	Quartic (4)	Quintic (5)	Nonlinear (6)
	STEM Track in Grade 11					
Comparative STEM Advantage	0.100***	0.005	0.039*	0.017	0.010	-0.065***
	(0.019)	(0.022)	(0.022)	(0.023)	(0.023)	(0.022)
Comparative STEM Advantage × Female	0.131***	0.179***	0.159***	0.159***	0.167***	0.155***
	(0.025)	(0.026)	(0.026)	(0.027)	(0.027)	(0.026)
Perc. Rank in STEM	0.136***	0.103***	0.099***	0.098***	0.105***	0.150***
	(0.026)	(0.026)	(0.026)	(0.027)	(0.027)	(0.026)
Perc. Rank non-STEM	-0.146***	-0.075***	-0.080***	-0.079***	-0.077***	-0.160***
	(0.022)	(0.024)	(0.024)	(0.024)	(0.024)	(0.023)
Perc. Rank in STEM × Female	0.031	0.040	0.035	0.031	0.022	0.028
	(0.034)	(0.033)	(0.034)	(0.035)	(0.035)	(0.034)
Perc. Rank non-STEM × Female	-0.129***	-0.098***	-0.121***	-0.121***	-0.117***	-0.110***
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.026)
Obs.	72,940	72,940	72,940	72,940	72,940	72,940
	Application for STEM University Degree					
Comparative STEM Advantage	0.003	-0.043	-0.043	-0.032	-0.030	-0.067**
	(0.026)	(0.027)	(0.027)	(0.029)	(0.029)	(0.029)
Comparative STEM Advantage × Female	0.139***	0.151***	0.139***	0.117***	0.120***	0.108***
	(0.035)	(0.035)	(0.035)	(0.036)	(0.036)	(0.036)
Perc. Rank in STEM	0.250***	0.071**	0.065*	0.056	0.053	0.090**
	(0.037)	(0.035)	(0.036)	(0.036)	(0.036)	(0.036)
Perc. Rank non-STEM	-0.225***	0.004	-0.000	0.011	0.010	-0.067**
	(0.031)	(0.031)	(0.031)	(0.032)	(0.032)	(0.031)
Perc. Rank in STEM × Female	-0.075	-0.072	-0.055	-0.011	-0.015	-0.011
	(0.048)	(0.046)	(0.047)	(0.049)	(0.049)	(0.049)
Perc. Rank non-STEM × Female	0.058	0.002	0.005	-0.023	-0.021	-0.007
	(0.037)	(0.037)	(0.037)	(0.038)	(0.038)	(0.038)
Obs.	45,259	45,259	45,259	45,259	45,259	45,259
Classroom FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated effects of comparative STEM advantage on track and university degree choices. The outcome variables are: (a) an indicator for whether a student enrolls in a STEM track in grade 11 (top panel), and (b) an indicator for whether a student applies for a STEM university degree two years later (bottom panel). For each of the two outcomes, we run different specifications for different degrees of polynomials for the absolute STEM advantage (columns (1)-(5)) and a nonlinear specification using binary indicators for each decile of absolute STEM advantage (column (6)). Each regression controls for student gender, absolute STEM advantage, STEM, non-STEM performance, percentile rank in STEM and non-STEM performance within the classroom, interactions of individual terms with gender, and classroom FE. Standard errors are clustered at the school cohort level. The last row in each panel shows the slope coefficient of the regression of each outcome variable on a female indicator, reflecting the gender gap in that outcome. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A7: **Effect of Comparative STEM Advantage on STEM Study Controlling for Ordinal Rank—Instead of Performance—in STEM and Non-STEM Subjects**

	Linear	Quadratic	Cubic	Quartic	Quintic	Nonlinear
	(1)	(2)	(3)	(4)	(5)	(6)
	STEM Track in Grade 11					
Comparative STEM Advantage	0.232***	0.080***	0.084***	0.010	0.004	0.005
	(0.019)	(0.021)	(0.022)	(0.022)	(0.022)	(0.022)
Comparative STEM Advantage × Female	0.103***	0.171***	0.124***	0.116***	0.121***	0.120***
	(0.024)	(0.025)	(0.026)	(0.025)	(0.025)	(0.025)
Obs.	72,940	72,940	72,940	72,940	72,940	72,940
Mean of Y	0.63	0.63	0.63	0.63	0.63	0.63
St. Dev. Y	0.48	0.48	0.48	0.48	0.48	0.48
Raw Gender Gap Y	-0.34	-0.34	-0.34	-0.34	-0.34	-0.34
	Application for STEM University Degree					
Comparative STEM Advantage	-0.042*	-0.143***	-0.135***	-0.114***	-0.108***	-0.093***
	(0.025)	(0.027)	(0.028)	(0.029)	(0.029)	(0.030)
Comparative STEM Advantage × Female	0.129***	0.148***	0.146***	0.143***	0.140***	0.134***
	(0.033)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
Obs.	45,259	45,259	45,259	45,259	45,259	45,259
Mean of Y	0.72	0.72	0.72	0.72	0.72	0.72
St. Dev. Y	0.45	0.45	0.45	0.45	0.45	0.45
Raw Gender Gap Y	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03
Classroom FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated effects of comparative STEM advantage on track and university degree choices. The outcome variables are: (a) an indicator for whether a student enrolls in a STEM track in grade 11 (top panel), and (b) an indicator for whether a student applies for a STEM university degree two years later (bottom panel). For each of the two outcomes, we run different specifications for different degrees of polynomials for the absolute STEM advantage (columns (1)-(5)) and a nonlinear specification using binary indicators for each decile of absolute STEM advantage (column (6)). Each regression controls for student gender, absolute STEM advantage, ordinal rank in STEM and non-STEM performance in the classroom, interactions of individual terms with gender, and classroom FE. Standard errors are clustered at the school cohort level. The last row in each panel shows the slope coefficient of the regression of each outcome variable on a female indicator, reflecting the gender gap in that outcome. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A8: **Effect of Comparative STEM Advantage on STEM Study Controlling for Performance in Each STEM and Non-STEM Subject**

	Linear	Quadratic	Cubic	Quartic	Quintic	Nonlinear
	(1)	(2)	(3)	(4)	(5)	(6)
	STEM Track in Grade 11					
Comparative STEM Advantage	0.138***	0.048**	0.074***	0.053**	0.048**	0.046**
	(0.018)	(0.020)	(0.020)	(0.021)	(0.021)	(0.021)
Comparative STEM Advantage × Female	0.136***	0.185***	0.150***	0.143***	0.146***	0.145***
	(0.021)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Obs.	72,937	72,937	72,937	72,937	72,937	72,937
Mean of Y	0.63	0.63	0.63	0.63	0.63	0.63
St. Dev. Y	0.48	0.48	0.48	0.48	0.48	0.48
Raw Gender Gap Y	-0.34	-0.34	-0.34	-0.34	-0.34	-0.34
	Application for STEM University Degree					
Comparative STEM Advantage	0.071***	-0.038	-0.043*	-0.037	-0.030	-0.011
	(0.024)	(0.026)	(0.026)	(0.027)	(0.027)	(0.028)
Comparative STEM Advantage × Female	0.087***	0.105***	0.106***	0.104***	0.101***	0.094***
	(0.027)	(0.028)	(0.028)	(0.029)	(0.029)	(0.029)
Obs.	45,258	45,258	45,258	45,258	45,258	45,258
Mean of Y	0.72	0.72	0.72	0.72	0.72	0.72
St. Dev. Y	0.45	0.45	0.45	0.45	0.45	0.45
Raw Gender Gap Y	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03
Classroom FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated effects of comparative STEM advantage on track and university degree choices. The outcome variables are: (a) an indicator for whether a student enrolls in a STEM track in grade 11 (top panel), and (b) an indicator for whether a student applies for a STEM university degree two years later (bottom panel). For each of the two outcomes, we run different specifications for different degrees of polynomials for the absolute STEM advantage (columns (1)-(5)) and a nonlinear specification using binary indicators for each decile of absolute STEM advantage (column (6)). Each regression controls for student gender, absolute STEM advantage, performance in eacg STEM and non-STEM subject, interactions of individual terms with gender, and classroom FE. Standard errors are clustered at the school cohort level. The last row in each panel shows the slope coefficient of the regression of each outcome variable on a female indicator, reflecting the gender gap in that outcome. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A9: **Balancing by Ordinal STEM Rank**

	Independent Variable: Ordinal STEM Rank
	(1)
Class Av. GPA	-0.00022
	(0.00037)
Class Median GPA	-0.00020
	(0.00053)
Prop. Female	-0.00005
	(0.00007)
Av. GPA Female	0.00032
	(0.00054)
Av. GPA Male	0.00032
	(0.00054)
Av. STEM GPA Female	-0.00061
	(0.00096)
Av. STEM GPA Male	0.00157
	(0.00100)
Av. non-STEM GPA Female	-0.00119
	(0.00076)
Av. non-STEM GPA Male	-0.00097
	(0.00083)

Notes: This table shows balancing checks of ordinal STEM rank with respect to the listed classroom characteristics. Each row represents a regression with the listed classroom characteristic as the dependent variable and individual ordinal rank in STEM advantage as the independent variable. All regressions control for school FE. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Dealing with Possible Sample Selection

In order to deal with possible bias caused by sample selection, we employ an Inverse-probability-weighted estimator for model (6). In particular, we assume that the process that causes some of the data to be missing is a function of observable covariates and a random process that is independent of the outcome. First, we formally test which observable students or class's characteristics are correlated with attrition, by estimating a probit model for attrition. Perhaps not surprisingly, attrition is correlated with student performance in STEM and non-STEM subjects, as well as students' STEM advantage. Additional variables that are significant predictors of attrition are the dummy for female, the interaction between non-STEM average performance and female, and average classroom performance in STEM and non-STEM subjects. Chi-square statistics for the Wald test of whether these variables are jointly equal to zero is 757.64, suggesting that these variables are jointly statistically different from zero at the highest level of significance. In other words, these variables are significant predictors of students' transfer out.

Given that the transfer-out rate may be non-random, we compute the inverse probability weights for Model (6) to correct for attrition. We compute the predicted probabilities and the inverse probability weights from the restricted probit. Intuitively, this procedure gives higher weight to students with characteristics similar to those of students who subsequently transfer out. Table A12 shows the main result for our model when attrition is

not controlled for and when it is controlled. The results remain largely unaffected.

Table A10: **Gender Difference in Early Leavers and Attrition Rate**

	Male	Female	Difference	*p-value*
	(1)	(2)	(3)	(4)
Early leavers	0.082	0.043	-0.039	0.000
Students' attrition	0.172	0.133	-0.039	0.000

Notes: This table reports male and female early leavers and attrition rate (in columns (1) and (2) respectively). Column (3) reports gender differences in early leavers and student attrition. Column (4) reports the *p*-value for the *t*-test on the gender difference.

Table A11: **Association between Classroom Performance and Gender Difference in Sample Attrition**

	GD Early Leavers	GD Students' Attrition
	(1)	(2)
Classroom GPA	-0.272	-0.029
	(0.195)	(0.223)
Obs.	3,428	3,428
School $\times$ Year FE	Yes	Yes

Notes: This table reports results of the estimated effects of the classroom average GPA on two types of attrition. Column (1) shows results of the estimated effect of classroom average GPA on the gender difference (GD, male minus female) in early leavers in each classroom. We define *early leavers* as those students who do not complete grade 10, but drop out from school early during their 10^{th} grade. Column (2) shows results of the estimated effect of classroom average GPA on the gender difference (GD) in students' attrition in each classroom. We define *attriters* as students who leave the sample at the end of 10^{th} grade after they complete grade 10. The unit of observation is the classroom. Clustered standard errors at the school level are reported in parentheses.

Table A12: **Effect of Comparative STEM Advantage on STEM track in Grade 11 without and with IPWs**

	Without Attrition Weights		With Attrition Weights	
	Quadratic	Nonlinear	Quadratic	Nonlinear
	(1)	(2)	(3)	(4)
Comparative STEM Advantage	0.038*	0.030	0.073***	0.047*
	(0.020)	(0.021)	(0.024)	(0.026)
Comparative STEM Advantage × Female	0.202***	0.161***	0.182***	0.139***
	(0.022)	(0.022)	(0.025)	(0.025)
Obs.	72,940	72,940	72,865	72,865
School × Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Mean Y	0.63	0.63	0.63	0.63
St. Dev Y	0.48	0.48	0.48	0.48

Notes: This table reports OLS estimates for model (6), without correcting for attrition (columns (1) and (2)), and using inverted probability weights to account for attrition (columns (3) and (4)). In each regression, the dependent variable is a dummy for whether the student applies to a STEM track at the end of grade 10. The first specification includes a quadratic term for STEM advantage and the second specification includes 10 dummies for different levels of absolute STEM advantage. Each regression controls for student gender, absolute STEM advantage, STEM, non-STEM performance, interactions of individual terms with gender, and classroom FE. Standard errors are clustered at the school cohort level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Performance in the First Semester of 10th Grade

As a robustness exercise, we use performance in the first semester of 10th grade to compute students' absolute and comparative STEM advantage. Students are allocated to classrooms at the beginning of grade 10. Therefore, a student's final exam scores at the end of the year, which determine their comparative STEM advantage, could be affected by peer effects. In our main analysis, this problem is mitigated by the fact that classroom average performance is controlled for through classroom FE. We use performance during the first semester in grade 10 as a robustness check. Table A13 shows the summary statistics when performance in the first semester of 10th grade is used. Figure A8 shows the performance in the first semester of 10th grade for males and females in Algebra, Physics, Chemistry, Modern Greek, Greek Literature, and Ancient Greek. Table 8 reports the estimates of our main model using first-semester performance. The results remain similar.

Table A13: **Descriptive Statistics of Performance in First Semester 10th Grade**

	Male	Female	Difference	*p-value*
	(1)	(2)	(3)	(4)
Panel A: Performance in Grade 10				
Algebra	14.078	14.556	0.478	0.000
Physics	14.277	14.591	0.314	0.000
Chemistry	14.594	15.144	0.550	0.000
Modern Greek	13.891	15.057	1.166	0.000
Greek Literature	14.378	15.807	1.429	0.000
Ancient Greek	13.891	15.214	1.323	0.000
Panel B: Constructed Variables in Grade 10				
Own Grade in STEM	14.315	14.763	0.448	0.000
Own Grade in non-STEM	14.052	15.357	1.305	0.000
Class Average Grade in STEM	14.541	14.565	0.024	0.001
Class Average Grade in non-STEM	14.754	14.761	0.007	0.346
Comparative STEM Advantage	0.456	0.316	-0.140	0.000

Notes: This table reports the gender differences in performance for the six subjects we use to construct our variable in grade 10 (Panel A) and the gender differences for the variable we construct and we use for our analysis (Panel B). The fourth column reports *p*-values for the *t*-test on the gender difference on each variables.

Table A14: **Effect of Comparative STEM Advantage on STEM Study Using Different Definitions of STEM Departments**

	Linear	Quadratic	Cubic	Quartic	Quintic	Nonlinear
	(1)	(2)	(3)	(4)	(5)	(6)
	Application for STEM University Degree					
Comparative STEM Advantage	0.072***	-0.041	-0.046*	-0.040	-0.034	-0.014
	(0.024)	(0.026)	(0.026)	(0.027)	(0.027)	(0.028)
Comparative STEM Advantage × Female	0.094***	0.113***	0.113***	0.112***	0.109***	0.102***
	(0.027)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Obs.	45,259	45,259	45,259	45,259	45,259	45,259
Mean of Y	0.72	0.72	0.72	0.72	0.72	0.72
St. Dev. Y	0.45	0.45	0.45	0.45	0.45	0.45
Raw Gender Gap Y	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03
	STEM departments = Science, Engineering, Technology, Economics, and Business					
Comparative STEM Advantage	0.114***	-0.005	0.001	-0.011	-0.011	0.008
	(0.021)	(0.022)	(0.023)	(0.024)	(0.024)	(0.024)
Comparative STEM Advantage × Female	0.128***	0.174***	0.153***	0.147***	0.147***	0.142***
	(0.023)	(0.024)	(0.025)	(0.025)	(0.025)	(0.025)
Obs.	72,940	72,940	72,940	72,940	72,940	72,940
Mean of Y	0.55	0.55	0.55	0.55	0.55	0.55
St. Dev. Y	0.50	0.50	0.50	0.50	0.50	0.50
Raw Gender Gap Y	-0.22	-0.22	-0.22	-0.22	-0.22	-0.22
	STEM departments = Science, Engineering, Technology, and Health Sciences					
Comparative STEM Advantage	0.133***	0.065***	0.056**	0.041	0.040	0.045*
	(0.022)	(0.024)	(0.024)	(0.025)	(0.025)	(0.025)
Comparative STEM Advantage × Female	0.110***	0.146***	0.127***	0.126***	0.125***	0.121***
	(0.023)	(0.024)	(0.025)	(0.025)	(0.025)	(0.025)
Obs.	72,940	72,940	72,940	72,940	72,940	72,940
Mean of Y	0.61	0.61	0.61	0.61	0.61	0.61
St. Dev. Y	0.49	0.49	0.49	0.49	0.49	0.49
Raw Gender Gap Y	-0.07	-0.07	-0.07	-0.07	-0.07	-0.07
Classroom FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports OLS estimates for model (6) for application to university STEM departments, using broader definitions of STEM departments. Panel A displays the effect using the same definition of STEM used in the main analysis for comparison purposes. Panel B displays the results when Economics and Business departments are included in the definition of STEM. Panel C shows the results when Health Science departments are included in the STEM definition. In each panel, we show several specifications for different degrees of polynomials for STEM advantage (columns (1)-(5)) as well as a nonlinear specification that uses dummy variables for each decile of absolute STEM advantage (column (6)). Each regression controls for student gender, absolute STEM advantage, STEM, non-STEM performance, interactions of individual terms with gender, and classroom FE. Standard errors are clustered at the school cohort level. The last row in each panel shows the slope coefficient of the regression of each outcome variable on a female indicator, reflecting the gender gap in that outcome. Each regression controls for student STEM performance, non-STEM performance, and absolute STEM advantage. Each regression includes classroom FE. Standard errors are clustered at the school cohort level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A15: **Effect of Comparative STEM Advantage on Science Track Choice in Grade 11**

	Linear	Quadratic	Cubic	Quartic	Quintic	Nonlinear
	(1)	(2)	(3)	(4)	(5)	(6)
	STEM Track in Grade 11					
Comparative STEM Advantage	-0.038**	0.086***	0.076***	0.063***	0.058***	0.019
	(0.019)	(0.020)	(0.020)	(0.021)	(0.021)	(0.021)
Comparative STEM Advantage × Female	0.099***	0.088***	0.060***	0.057***	0.059***	0.064***
	(0.020)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Obs.	72,940	72,940	72,940	72,940	72,940	72,940
Mean of Y	0.20	0.20	0.20	0.20	0.20	0.20
St. Dev. Y	0.40	0.40	0.40	0.40	0.40	0.40
School × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated effects of comparative STEM advantage on choosing the Science track in grade 11. The outcome variables is an indicator capturing whether a student enrolls in a Science track in grade 11. We run different specifications for different degrees of polynomials for the absolute STEM advantage (columns (1)-(5)) and a nonlinear specification using binary indicators for each decile of absolute STEM advantage (column (6)). Each regression controls for student gender, absolute STEM advantage, STEM, non-STEM performance, interactions of individual terms with gender, and school-year FE. Standard errors are clustered at the school cohort level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A16: **Effect of Comparative STEM Advantage Excluding Students who Apply to a Degree Program in Business or Economics**

	Linear	Quadratic	Cubic	Quartic	Quintic	Nonlinear
	(1)	(2)	(3)	(4)	(5)	(6)
	STEM Track in Grade 11					
Comparative STEM Advantage	0.134***	0.076***	0.092***	0.067***	0.059**	0.049*
	(0.021)	(0.025)	(0.025)	(0.026)	(0.026)	(0.026)
Comparative STEM Advantage × Female	0.133***	0.192***	0.140***	0.139***	0.142***	0.140***
	(0.024)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
Obs.	52,901	52,901	52,901	52,901	52,901	52,901
Mean of Y	0.56	0.56	0.56	0.56	0.56	0.56
St. Dev. Y	0.50	0.50	0.50	0.50	0.50	0.50
Raw Gender Gap Y	-0.34	-0.34	-0.34	-0.34	-0.34	-0.34
	Application for STEM University Degree					
Comparative STEM Advantage	0.077***	-0.052	-0.059*	-0.051	-0.041	-0.018
	(0.029)	(0.032)	(0.033)	(0.034)	(0.034)	(0.035)
Comparative STEM Advantage × Female	0.079**	0.086**	0.089**	0.083**	0.078**	0.066*
	(0.035)	(0.037)	(0.037)	(0.037)	(0.037)	(0.037)
Obs.	29,441	29,441	29,441	29,441	29,441	29,441
Mean of Y	0.67	0.67	0.67	0.67	0.67	0.67
St. Dev. Y	0.47	0.47	0.47	0.47	0.47	0.47
Raw Gender Gap Y	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03
School × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated effects of comparative STEM advantage on track and university degree choices, excluding students who apply to a university degree program in Business or Economics. The outcome variables are: (a) an indicator for whether a student enrolls in a STEM track in grade 11 (top panel), and (b) an indicator for whether a student applies for a STEM university degree two years later (bottom panel). For each of the two outcomes, we run different specifications for different degrees of polynomials for the absolute STEM advantage (columns (1)-(5)) and a nonlinear specification using binary indicators for each decile of absolute STEM advantage (column (6)). Each regression controls for student gender, absolute STEM advantage, STEM, non-STEM performance, interactions of individual terms with gender, and classroom FE. Standard errors are clustered at the school cohort level. The last row in each panel shows the slope coefficient of the regression of each outcome variable on a female indicator, reflecting the gender gap in that outcome. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A17: **Effect of Comparative STEM Advantage on University Degree Choices**

	Linear (1)	Quadratic (2)	Cubic (3)	Quartic (4)	Quintic (5)	Nonlinear (6)
	Application for Engineering and Technology University Degree					
Comparative STEM Advantage	0.085***	-0.044	-0.046*	-0.032	-0.026	0.000
	(0.026)	(0.028)	(0.028)	(0.029)	(0.029)	(0.030)
Comparative STEM Advantage × Female	0.106***	0.120***	0.124***	0.123***	0.121***	0.111***
	(0.031)	(0.032)	(0.032)	(0.033)	(0.033)	(0.033)
Obs.	40,789	40,789	40,789	40,789	40,789	40,789
Mean of Y	0.66	0.66	0.66	0.66	0.66	0.66
St. Dev. Y	0.48	0.48	0.48	0.48	0.48	0.48
Raw Gender Gap Y						
	Application for Science and Math University Degree					
Comparative STEM Advantage	0.038	-0.044	-0.055	-0.065*	-0.056	-0.041
	(0.030)	(0.033)	(0.034)	(0.036)	(0.037)	(0.037)
Comparative STEM Advantage × Female	0.109***	0.118***	0.120***	0.122***	0.117***	0.109***
	(0.035)	(0.036)	(0.036)	(0.037)	(0.037)	(0.037)
Obs.	28,970	28,970	28,970	28,970	28,970	28,970
Mean of Y	0.51	0.51	0.51	0.51	0.51	0.51
St. Dev. Y	0.50	0.50	0.50	0.50	0.50	0.50
Raw Gender Gap Y						
Classroom FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated effects of comparative STEM advantage on track and university degree choices. The outcome variables are: (a) an indicator for whether a student enrolls in a Engineering or Technology university degree (top panel), and (b) an indicator for whether a student applies for a Mathematics and Science university degree two years later (bottom panel). For each of the two outcomes, we run different specifications for different degrees of polynomials for the absolute STEM advantage (columns (1)-(5)) and a nonlinear specification using binary indicators for each decile of absolute STEM advantage (column (6)). Each regression controls for student gender, absolute STEM advantage, STEM, non-STEM performance, interactions of individual terms with gender, and classroom FE. Standard errors are clustered at the school cohort level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A18: **Effect of Comparative STEM Advantage on STEM Study Using Different Definitions of STEM Subjects**

	STEM Track in Grade 11			
	(1)	(2)	(3)	(4)
Comparative STEM Advantage	0.030			
	(0.022)			
Comparative STEM Advantage × Female	0.160***			
	(0.023)			
Comparative STEM Advantage (STEM=Algebra)		0.045**		
		(0.021)		
Comparative STEM Advantage (STEM=Algebra) × Female		0.152***		
		(0.023)		
Comparative STEM Advantage (STEM=Chemistry)			0.050**	
			(0.021)	
Comparative STEM Advantage (STEM=Chemistry) × Female			0.150***	
			(0.021)	
Comparative STEM Advantage (STEM=Physics)				0.050**
				(0.021)
Comparative STEM Advantage (STEM=Physics) × Female				0.110***
				(0.022)
Obs.	72,940	72,940	72,940	72,940
Classroom FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Notes: This table reports OLS estimates for model (6), using different definitions of STEM advantage. In column (1) STEM is defined as the average of Algebra, Chemistry, and Physics; in column (2) STEM is defined as performance only in Algebra; in column (3) STEM is defined as performance only in Chemistry; in column (4) STEM is defined as performance only in Physics. The non-STEM subjects average performance is always defined as average performance in Modern Greek, Greek Literature, and Ancient Greek. In each regression the dependent variable is a dummy indicating whether the student applied to a STEM track at the end of grade 10. Each regression controls for student gender, a second-order polynomial of absolute STEM advantage, STEM, non-STEM performance, interactions of individual terms with gender, and classroom FE. Standard errors are clustered at the school cohort level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A19: **Effect of Comparative STEM Advantage on STEM Study Controlling for the Number of Applications**

	Linear	Quadratic	Cubic	Quartic	Quintic	Nonlinear
	(1)	(2)	(3)	(4)	(5)	(6)
	Application for STEM University Degree					
Comparative STEM Advantage	0.072***	-0.041	-0.046*	-0.040	-0.034	-0.014
	(0.024)	(0.026)	(0.026)	(0.027)	(0.027)	(0.028)
Comparative STEM Advantage × Female	0.094***	0.113***	0.113***	0.112***	0.109***	0.102***
	(0.027)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
	Application for STEM University Degree controlling for the Number of University Degree Applications					
Comparative STEM Advantage	-0.043**	-0.013	-0.019	-0.022	-0.021	-0.025
	(0.017)	(0.019)	(0.019)	(0.021)	(0.021)	(0.021)
Comparative STEM Advantage × Female	0.076***	0.075***	0.071***	0.074***	0.074***	0.073***
	(0.022)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Obs.	45,259	45,259	45,259	45,259	45,259	45,259
Mean of Y	0.72	0.72	0.72	0.72	0.72	0.72
St. Dev. Y	0.45	0.45	0.45	0.45	0.45	0.45
Raw Gender Gap Y	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03
Classroom FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table (Panel B) reports OLS estimates for model (6) for application to university STEM departments when we control for the number of degree program students included in their preference list, which is equivalent to the number of degree program applications. The average student includes roughly 25 degree programs in their preference list with a standard deviation of 22.2. Panel A displays the results from the main analysis for comparison purposes. We show estimates from several specifications using different degrees of polynomials for STEM advantage (columns (1)-(5)) as well as a nonlinear specification that employs dummy variables for each decile of absolute STEM advantage (column (6)). Each regression controls for student gender, absolute STEM advantage, STEM, non-STEM performance, interactions of individual terms with gender, and classroom FE. Standard errors are clustered at the school cohort level. The last row shows the slope coefficient of the regression of each outcome variable on a female indicator, reflecting the gender gap in the outcome. Each regression controls for student STEM performance, non-STEM performance, and absolute STEM advantage. Each regression includes classroom FE. Standard errors are clustered at the school cohort level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

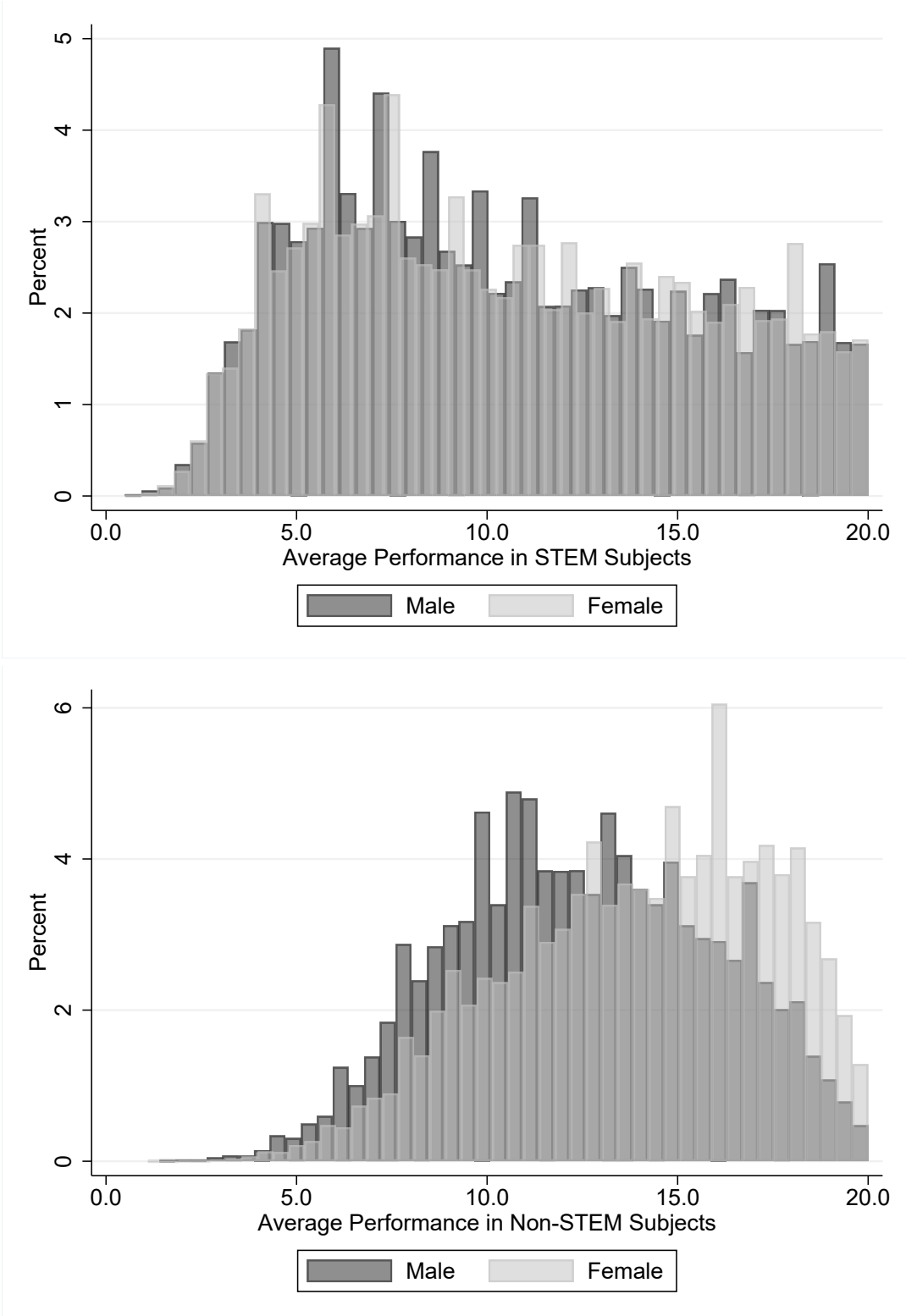
Table A20: **Effect of Comparative Non-STEM Advantage on STEM Study Outcomes**

	STEM Track in Grade 11		Applied for STEM University Degree	
	Quadratic	Nonlinear	Quadratic	Nonlinear
	(1)	(2)	(3)	(4)
Comparative non-STEM Adv.	-0.025	-0.035	-0.044	-0.014
	(0.021)	(0.023)	(0.028)	(0.031)
Comparative non-STEM Adv. × Female	-0.167***	-0.112***	-0.057	-0.042
	(0.030)	(0.034)	(0.045)	(0.048)
Obs.	72,940	72,940	45,259	45,259
Classroom FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Mean Y	0.63	0.63	0.72	0.72
St. Dev Y	0.48	0.48	0.45	0.45
Raw Gender Gap Y	-0.34	-0.34	-0.03	-0.03

Notes: This table reports the OLS estimates for model (6). Rank in non-STEM advantage is used rather than rank in STEM advantage. For each of the two outcomes (grade 11 STEM track choice and application to STEM degree program), two specifications are considered. Columns (1) and (3) show the effect of comparative non-STEM advantage, while columns (2) and (4) report the interaction term between comparative non-STEM advantage and the dummy for female. Each regression controls for student gender, absolute STEM advantage, STEM, non-STEM performance, interactions of individual terms with gender, and classroom FE. Standard errors are clustered at the school cohort level. The last row in each panel shows the slope coefficient of the regression of each outcome variable on a female indicator, reflecting the gender gap in that outcome. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

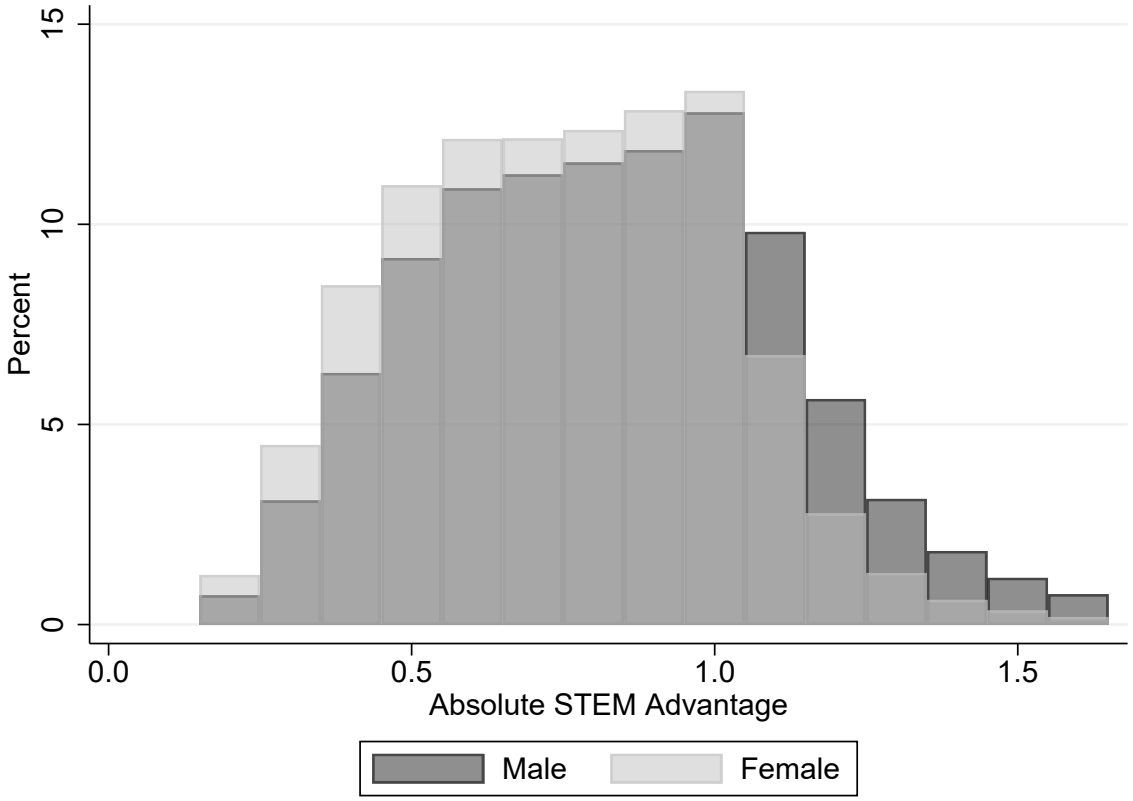
Appendix Figures

Figure A1: **Distribution of Performance in STEM and Non-STEM Subjects at the End of 10th Grade**



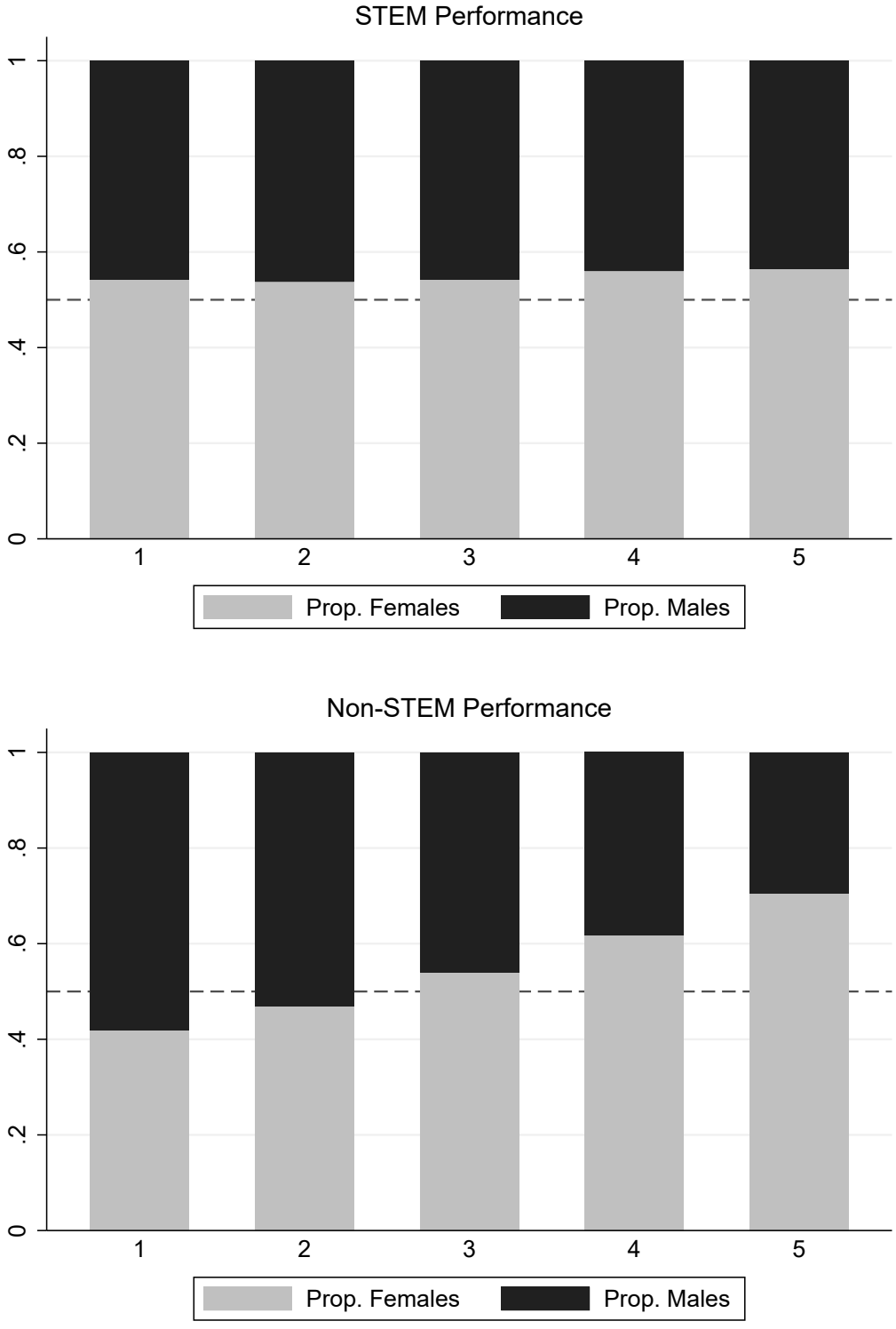
Notes: These two graphs plot the distributions of performance at the end of 10th grade for STEM subjects (Mathematics, Physics, and Chemistry) in the first graph and non-STEM subjects (Modern Greek, Ancient Greek, and Greek Literature) in the second graph.

Figure A2: **Distribution of Absolute STEM Advantage at the End of 10th Grade**



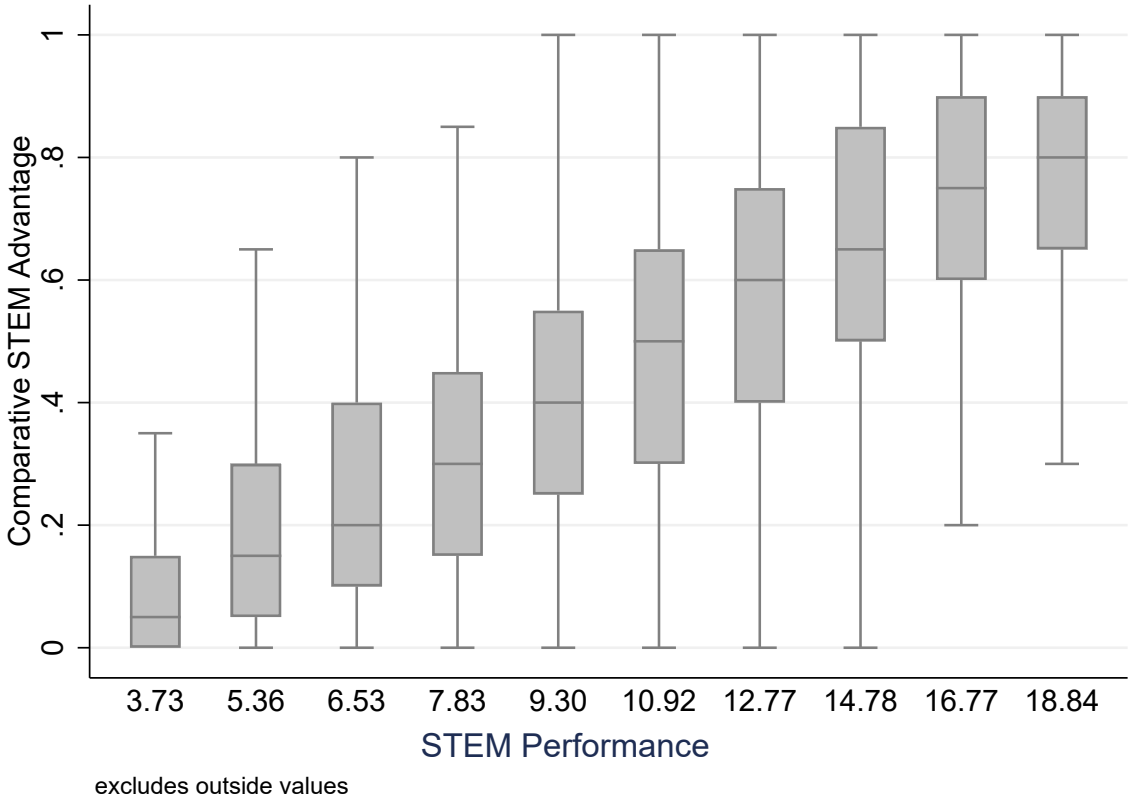
Notes: This graph plots the distribution of absolute STEM advantage at the end of grade 10 for males and females.

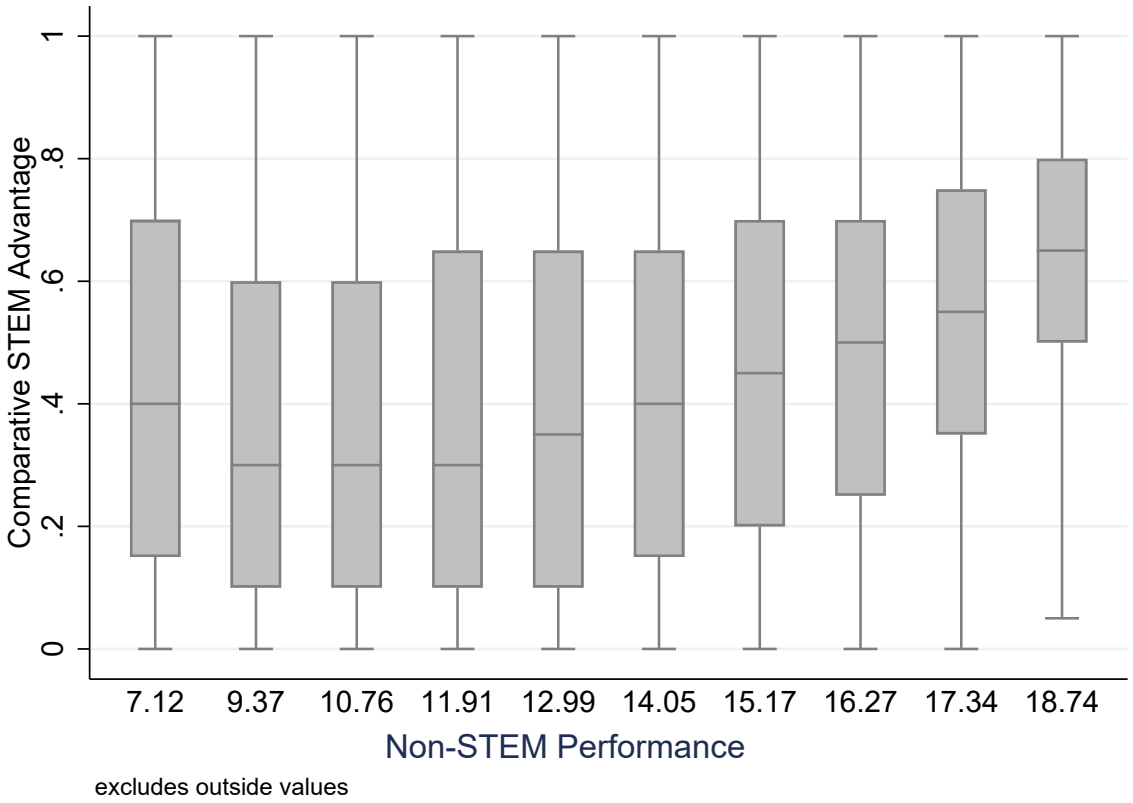
Figure A3: **Proportion of Males and Females by Quintile of STEM/Non-STEM Performance Distribution**



Notes: These two figures show the proportion of students by quintiles of the STEM and non-STEM performance distribution at the end of grade 10. STEM ability is computed as average GPA in Algebra, Physics, and Chemistry. Non-STEM performance is computed as average GPA in Modern Greek, Greek Literature, and Ancient Greek. While the proportion of females is constant across the quintile of STEM performance distribution, a greater number of female are in the top quintiles of the non-STEM performance distribution. The dotted red line is drawn at 0.5 to show the equal representation of gender as benchmark.

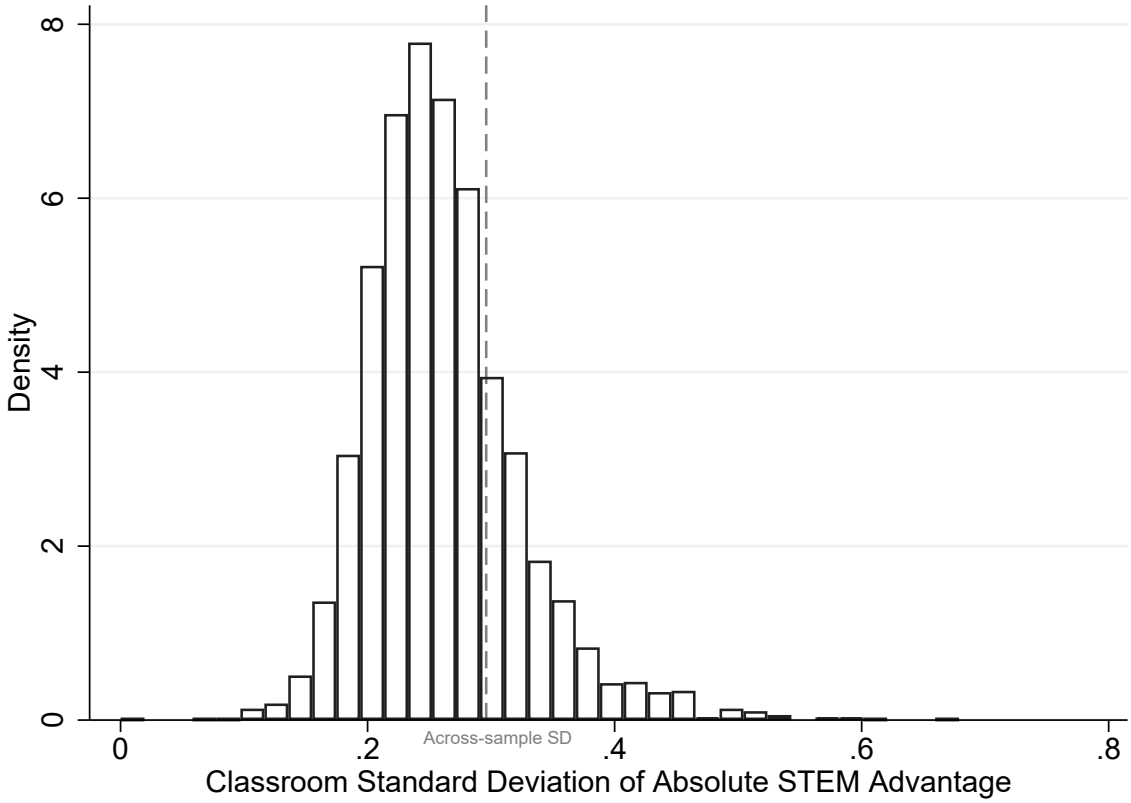
Figure A4: **Variation of Comparative STEM Advantage with Respect to STEM and Non-STEM Performance**





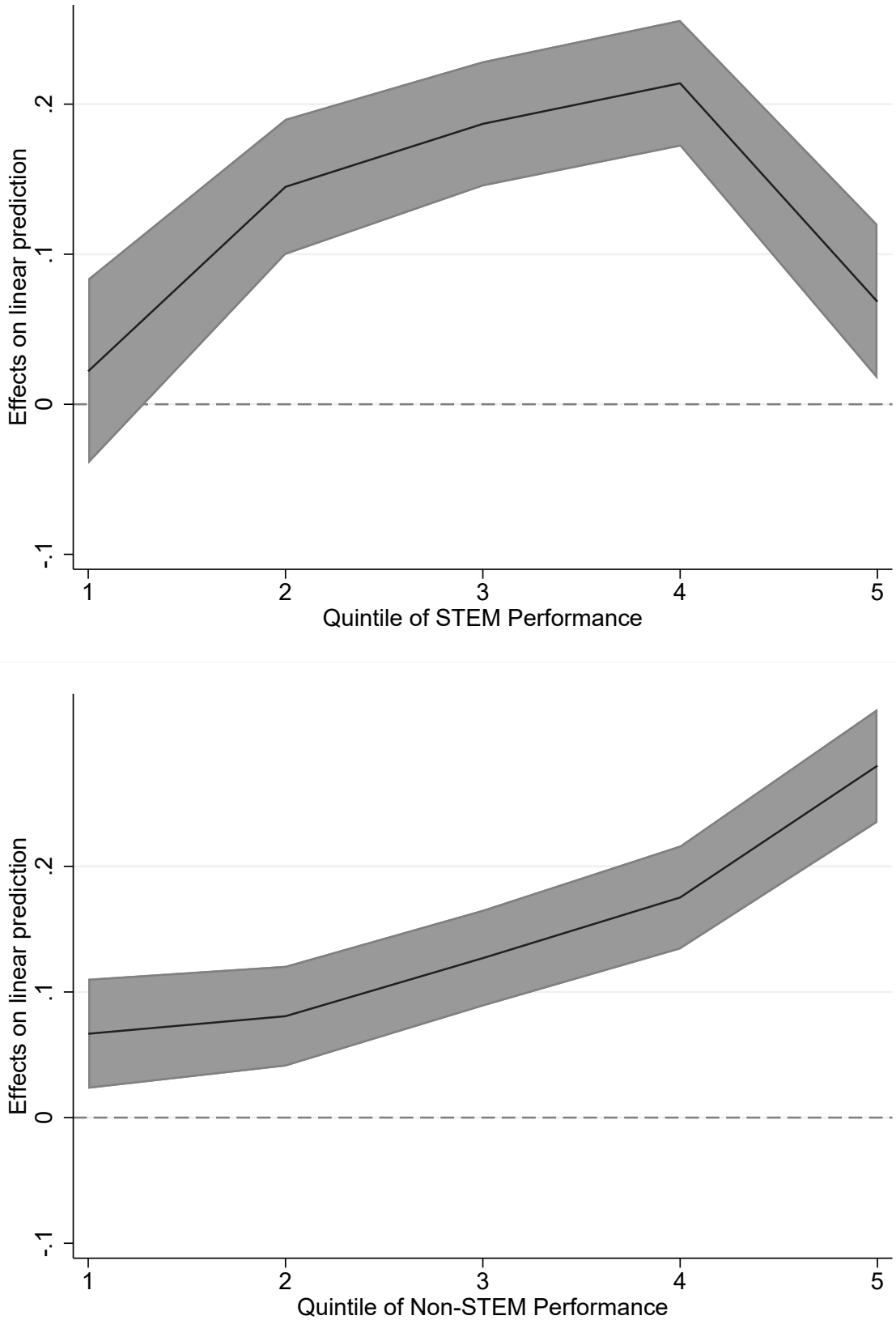
Notes: These two box plots show the variation in rank in STEM advantage by decile of STEM and non-STEM performance at the end of grade 10.

Figure A5: **Distribution of Dispersion of Absolute STEM Advantage within Classrooms**



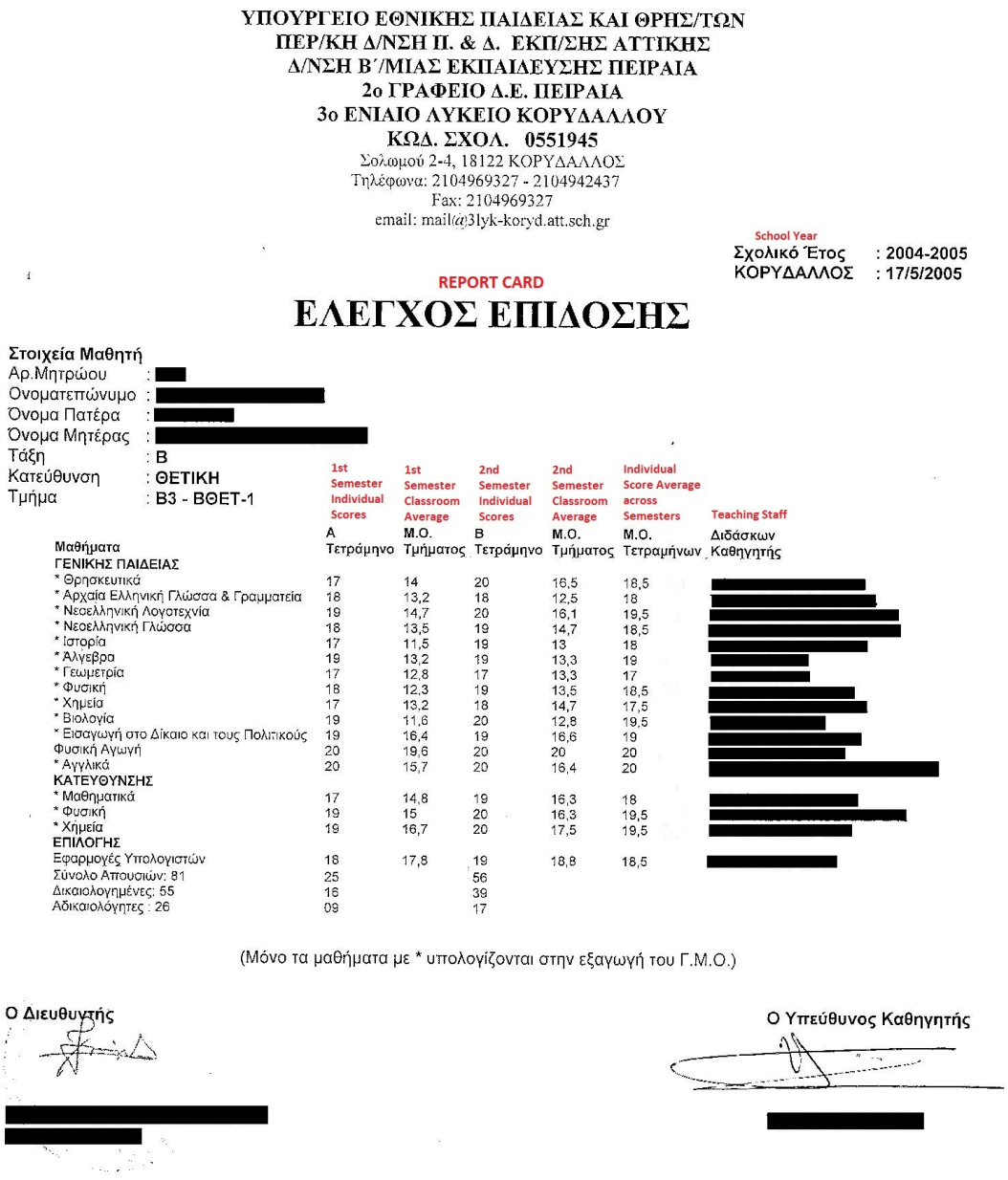
Notes: The histogram of within-classroom standard deviation of absolute STEM advantage reveals substantial variation in the dispersion of absolute STEM advantage in the classroom. The vertical line corresponds to the standard deviation of absolute STEM advantage across all students.

Figure A6: **Differential Effect of Comparative STEM Advantage across Different Quintiles STEM and Non-STEM Performance**



Notes: These two graphs plot the estimates for rank in STEM advantage as in model (6), on STEM track choice in grade 11, for each quintile of the STEM and non-STEM ability distribution. Both models include a quadratic polynomial for absolute STEM advantage. Standard errors are clustered at the school cohort level.

Figure A7: **Example of a High School Report Card**



ΥΠΟΥΡΓΕΙΟ ΕΘΝΙΚΗΣ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣ/ΤΩΝ
ΠΕΡ/ΚΗ Δ/ΝΣΗ Π. & Δ. ΕΚΠ/ΣΗΣ ΑΤΤΙΚΗΣ
Δ/ΝΣΗ Β΄/ΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ ΠΕΙΡΑΙΑ
2ο ΓΡΑΦΕΙΟ Δ.Ε. ΠΕΙΡΑΙΑ
3ο ΕΝΙΑΙΟ ΛΥΚΕΙΟ ΚΟΡΥΔΑΛΛΟΥ
ΚΩΔ. ΣΧΟΛ. 0551945
Σολωμού 2-4, 18122 ΚΟΡΥΔΑΛΛΟΣ
Τηλέφωνα: 2104969327 - 2104942437
Fax: 2104969327
email: mail@3lyk-koryd.att.sch.gr

School Year
Σχολικό Έτος : 2004-2005
ΚΟΡΥΔΑΛΛΟΣ : 17/5/2005

REPORT CARD

ΕΛΕΓΧΟΣ ΕΠΙΔΟΣΗΣ

Στοιχεία Μαθητή
Αρ.Μητρώου : [redacted]
Ονοματεπώνυμο : [redacted]
Όνομα Πατέρα : [redacted]
Όνομα Μητέρας : [redacted]
Τάξη : Β
Κατεύθυνση : ΘΕΤΙΚΗ
Τμήμα : Β3 - ΒΘΕΤ-1

Μαθήματα	1st Semester Individual Scores A Τετράμηνο	1st Semester Classroom Average Μ.Ο. Τμήματος	2nd Semester Individual Scores Β Τετράμηνο	2nd Semester Classroom Average Μ.Ο. Τμήματος	Individual Score Average across Semesters Μ.Ο. Τετραμήνων	Teaching Staff Διδάσκων Καθηγητής
ΓΕΝΙΚΗΣ ΠΑΙΔΕΙΑΣ						
* Θρησκευτικά	17	14	20	16,5	18,5	[redacted]
* Αρχαία Ελληνική Γλώσσα & Γραμματεία	18	13,2	18	12,5	18	[redacted]
* Νεοελληνική Λογοτεχνία	19	14,7	20	16,1	19,5	[redacted]
* Νεοελληνική Γλώσσα	18	13,5	19	14,7	18,5	[redacted]
* Ιστορία	17	11,5	19	13	18	[redacted]
* Άλγεβρα	19	13,2	19	13,3	19	[redacted]
* Γεωμετρία	17	12,8	17	13,3	17	[redacted]
* Φυσική	18	12,3	19	13,5	18,5	[redacted]
* Χημεία	17	13,2	18	14,7	17,5	[redacted]
* Βιολογία	19	11,6	20	12,8	19,5	[redacted]
* Εισαγωγή στο Δίκαιο και τους Πολιτικούς	19	16,4	19	16,6	19	[redacted]
Φυσική Αγωγή	20	19,6	20	20	20	[redacted]
* Αγγλικά	20	15,7	20	16,4	20	[redacted]
ΚΑΤΕΥΘΥΝΣΗΣ						
* Μαθηματικά	17	14,8	19	16,3	18	[redacted]
* Φυσική	19	15	20	16,3	19,5	[redacted]
* Χημεία	19	16,7	20	17,5	19,5	[redacted]
ΕΠΙΛΟΓΗΣ						
Εφαρμογές Υπολογιστών	18	17,8	19	18,8	18,5	[redacted]
Σύνολο Απουσιών: 81	25		56			
Δικαιολογημένες: 55	16		39			
Αδικαιολόγητες : 26	09		17			

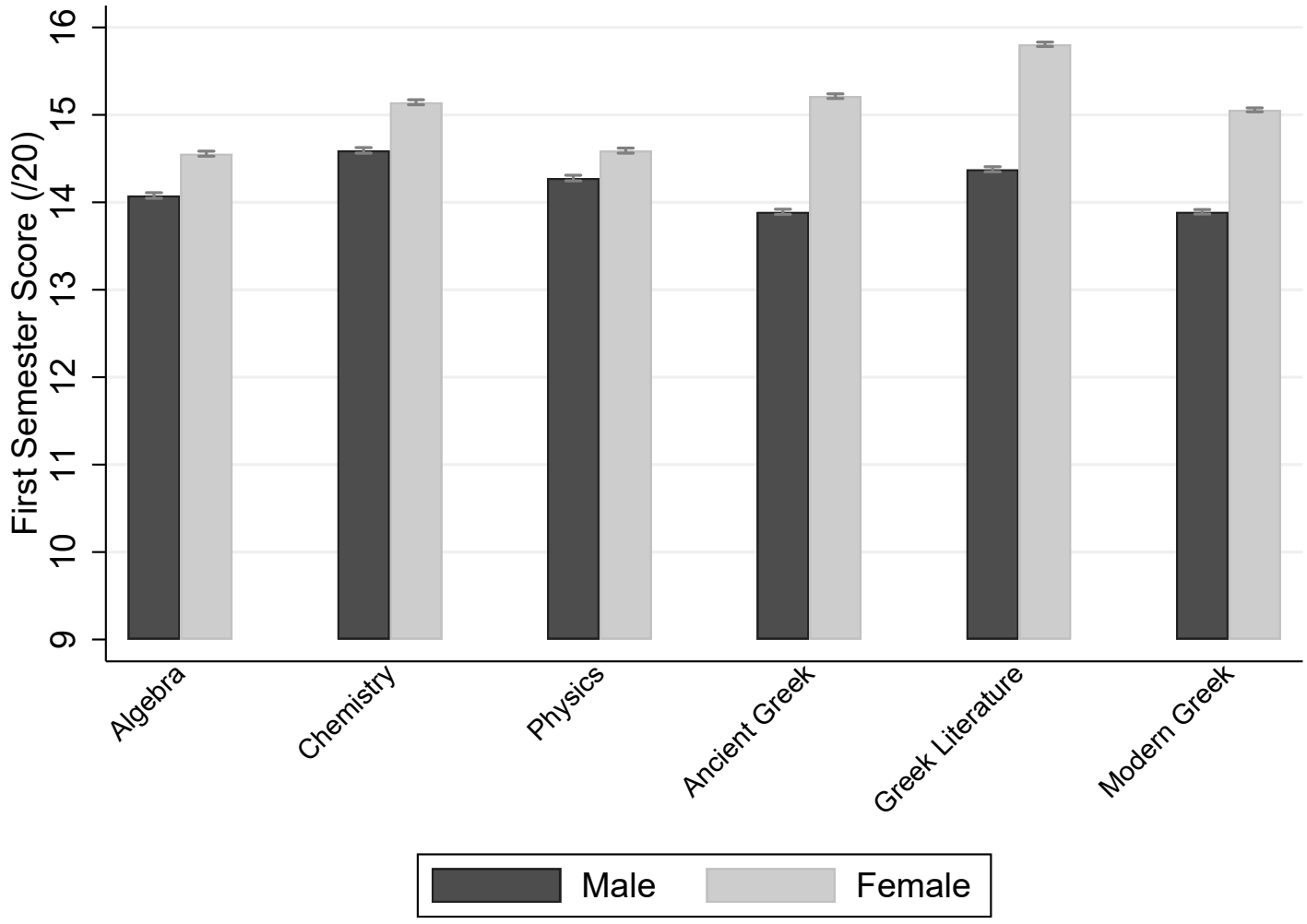
(Μόνο τα μαθήματα με * υπολογίζονται στην εξαγωγή του Γ.Μ.Ο.)

Ο Διευθυντής
[redacted]

Ο Υπεύθυνος Καθηγητής
[redacted]

Notes: The first column shows the subject taught. The second and fourth columns show individual scores in each subject in the first and second semester, respectively. The third and fifth columns show the classroom average score in each subject in the first and second semester, respectively. The sixth column shows the average individual scores in each subject across the two semesters. Scores range between 0 and 20 with 20 representing the highest performance. No score scaling is implemented. The seventh column shows the name of the teacher teaching each subject. At the end of the report card, students can see the number of their excused and unexcused absences in each semester. All names and student identifying information has been redacted.

Figure A8: **Performance in STEM and Non-STEM Subjects in the 1st Semester of 10th Grade by Gender**



Notes: The graph displays the performance in six subjects for males and females in the first semester of 10th grade along with the respective 95% confidence interval. Females perform significantly better in almost every subject, but their advantage is higher in non-STEM subjects (Modern Greek, Greek Literature, and Ancient Greek).