Heterogeneous Effects of the Gender Composition of Classrooms: Evidence from a Natural Experiment in Switzerland

Damiano Pregaldini* Uschi Backes-Gellner † Gerald Eisenkopf †

January 14, 2018

Preliminary draft. Please don't cite or circulate

Abstract

This paper investigates how gender composition of classes impacts achievements of students who selected into different specialization tracks based on individual preferences. Based on administrative records from one of the largest high schools in the canton of Zurich (Switzerland), we are able to identify the causal effect of the gender composition of classes on students' achievements by exploiting random assignment of students to classes as well as variation in gender composition across cohorts. Compared to the previous literature, which mainly focused on average effects, we find highly heterogeneous effects across students who selected into different specialization tracks. While girls and boys with a preference for languages tend to benefit from a higher proportion of girls in the class, the effect is negative for girls who choose the more mathematics intensive specialization track. This findings have important implications for the optimal organization of classes in schools, for the discussion on the costs and benefits of single-sex versus coeducational schooling, as well as for the explanation of career trajectories after school.

Keywords:

JEL Codes:

*Department of Business Administration, University of Zürich

†University of Vechta

1 Introduction

The effect of classroom composition on educational production has received major attention in personnel and education economics in the last decades (e.g., Lazear, 2001; Eisenkopf et al., 2015; Lavy and Schlosser, 2011; Hoxby, 2000; Oosterbeek and Van Ewijk, 2014). So far, the literature has analysed the relationship between gender and students' achievements from different perspectives. However, evidence on how this relationship varies across students with different characteristics and preferences is still scarce.

One strand of previous literature has focused on the dichotomy between coeducation and single-sex classes and has found mixed results. For example, Dustmann et al. (2017) and Park et al. (2013) find that students in single-sex classes outperform students in mixed gender classes using random assignment of students into single-sex versus coeducational high schools in Korea. Analysing data from a pedagogical college in Switzerland, Eisenkopf et al. (2015) find a positive effect on grades in mathematics only for girls. Jackson (2012), on the other hand, finds a statistically significant effect of being in a single-sex class only for girls with strong preferences for single-sex classes, suggesting the magnitude of the effect might depend on individual preferences.

Another strand of previous literature analysed how different proportions of girls in the class affect students' achievement. Lavy and Schlosser (2011) and Hoxby (2000) find that an increase in the proportion of girls in the classroom is associated with increased students' achievement, both for boys and for girls. Oosterbeek and Van Ewijk (2014) investigate the impact of the proportion of girls in universities on the dropout probability. They find that male students in classes with a relatively higher proportion of female students tend to postpone their decision to drop out. However, the results of these studies are limited to distributional implications, since the effect is found to be positive for both girls and boys. In fact, the gains from increasing the proportion of girls in a class is offset by the losses in the other classes. It is therefore important to better understand how the gender composition fo the class impacts achievements of different students.

This paper tries to close these gaps and contributes to the existing literature by analysing how the gender composition of the class affects achievements of students who selected into different specialization tracks at high school (Languages vs. Mathematics/Sciences). Because individuals select into different environments based on social preferences (Lazear et al., 2012), we can expect to observe different effects of classroom composition depending on these differences in the revealed social preferences. While we don't directly observe the determinants of selection, we expect students to sort into the more mathematics intensive specialization track according to their ability in mathematics, their self-confidence in the subject and, as pointed out by previous studies, based their competitiveness (Buser et al., 2014, 2017).

In our analysis, we exploit exogenous variation in gender composition of classrooms generated by a natural experiment in one of the largest high schools in the canton of Zurich. This randomized variation in gender composition allows us to causally identify the effect of the proportion of girls in the class on achievements of both girls and boys in mathematics and German. While most of the studies consider the average effect of gender composition on students' achievements, we analyse how this effect differs across students who self-selected into different specialization tracks (Maths/Sciences or Languages).

Understanding how achievements of boys and girls in different specialization tracks react to changes in the gender composition of the class allows for an optimal allocation of students to classes, thus enhancing school efficiency. Furthermore, the relationship is important to explain gender segregation in employment. Previous studies claim that girls perform better in mathematics when they are in girls dominated classes. If this is the case, and since school grades are likely to determine subsequent educational choices, changing the gender composition of classes might bring more female students into STEM [1] fields of study.

In line with the previous literature, our results indicate that, on average, a higher proportion of girls in the class is associated with higher grades in Mathematics for both girls and boys. On the contrary, we find no effect on grades in German. Our results show that effects are highly heterogeneous across specialization tracks. While a higher proportion of girls in the class positively affects achievements of both girls and boys in the specialization track Modern Languages, in the specialization track Mathematics/Sciences it has no effect on boys and a negative effect on girls.

2 Institutional background

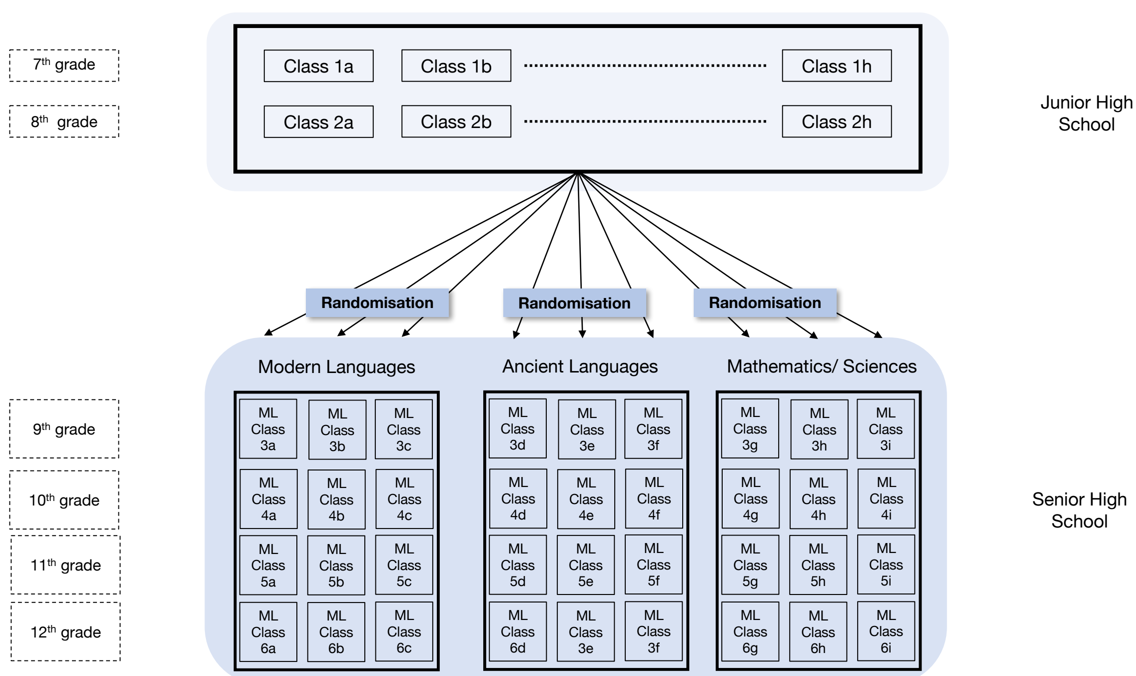
We analyse achievements of students in one of the largest schools of secondary education (high school) in the canton of Zurich. The school has a total of six grades divided between junior (grades 7th to 8th) and senior high schools (grades 9th to 12th) (cf. figure 1). Students typically enter the school in the 7th grade (12-13 years old), and attend junior and senior high schools, or in the 9th grade (14-15 years old) and attend only senior high school. Students usually graduate from high school at age 18-19.

While in junior high school all students follow the same course program, in senior high school, students can choose one of three specialization tracks among Modern Languages, Ancient Languages, and Maths/Sciences. Beside the common core curriculum, students in each specialization track takes specialized courses in English, Latin, Spanish, Italian or Russian (if they choose the specialization track Modern Languages), Latin and ancient Greek (if they choose the specialization track Ancient Languages), and biology, chemistry, application of mathematics or physics (if they choose the specialization track Maths/Sciences).

[1]Science, technology, engineering and mathematics

Students who were in the same class during during junior high school might choose different specialization tracks for senior high school. Therefore, a reorganisation of the classes is necessary in the transition from junior to senior high school. To avoid any favouritism and parent politicking, the administration of the school applies a strict random assignment of students to classes within the chosen specialization track, as illustrated in figure 1. At the same time, also the assignment of teachers to classes within specialization tracks doesn't follow a clear pattern. Even though some teachers are more prone to teach classes in a specific specialization track, within specialization tracks teachers has less freedom to choose a specific class.

Figure 1: Structure of the high school



Note: The number of classes can vary depending on the cohort and on the specialization track.

3 Data

Our empirical analysis bases on administrative records on all students from grade 7th to 12th who were enrolled in the aforementioned high school in the canton of Zurich between autumn 2002 and spring 2012. The record contains students' semester grades, gender and gender composition of the class. The data also includes class and teacher identifiers, as well as the teacher's gender. Table 1 reports the total number of students and classes (in all 6 grades) in each year, as well as the number of new 7th grade classes and students [2]. Every year, there were between 154 and 216 new students entering the school in the 7th grade. The number of total students increased over time and reached the maximum of 1108 students in the school year 2010/2011. Because some students may have had to repeat terms, leave the high school because of insufficient academic performance, have moved away or joined the high school at a later stage, our dataset takes the form of an unbalanced panel [3]. We

Table 1: Structure of the data

	Total		New entries	
School year	Number of students	Classes	Number of students	Classes
2002/2003	929	44	176	8
2003/2004	969	46	154	7
2004/2005	962	46	177	7
2005/2006	976	47	166	8
2006/2007	975	48	177	8
2007/2008	990	48	167	8
2008/2009	1020	49	204	9
2009/2010	1078	51	215	10
2010/2011	1108	54	216	10
2011/2012	1096	54	177	10

restrict the sample to students in 9th grade and later, since randomization takes place in the transition from grade 8th to grade 9th. After accounting for missing values, the restricted sample contains 2269 students: 816 students in the specialization track Modern Languages,

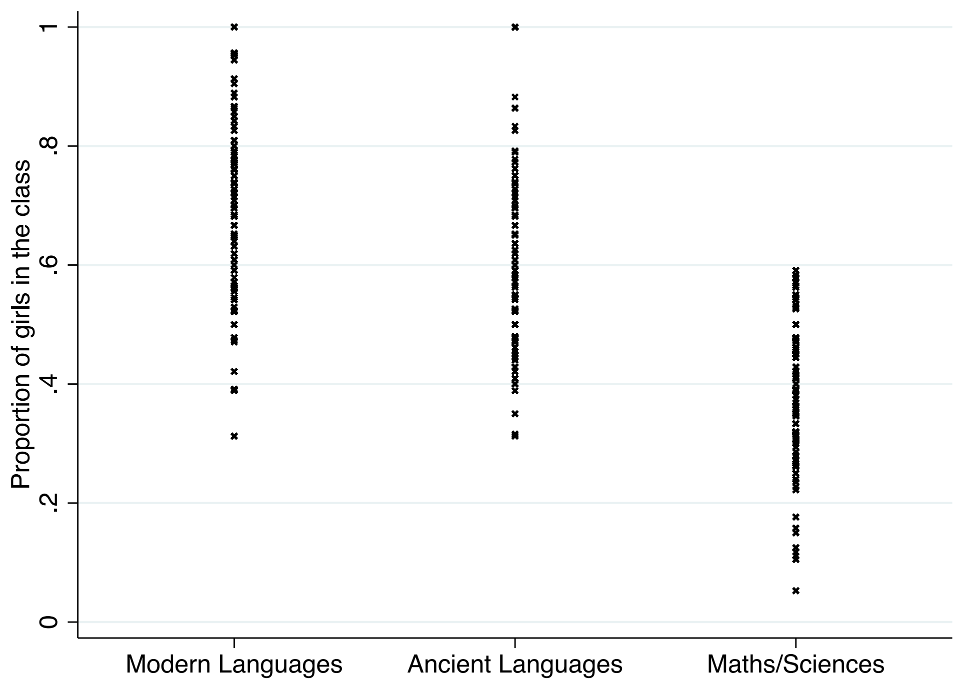
[2] As an example, in the school year 2002/2003 we have data on 929 students and 44 classes. The number of new students in that year was 176 divided in 8 classes.

[3] The share of students, for which we have at least one data point, who graduated from the high school is 83%

574 in Ancient Languages and 879 in Maths/Sciences.

The explanatory variable of interest is the gender composition of the class, which we operationalize as the proportion of girls in the class. This variable is constructed by dividing the number of girls in a given class by the total number of students in that class. Not surprisingly, gender composition of classes varies strongly between specialization tracks, as illustrated in figure 2. In fact, girls are particularly well represented in the specialization tracks Modern and Ancient Languages (proportion of girls between 31 and 100%). On the contrary, classes in the specialization track Maths/Sciences feature, on average, a lower proportion of girls (proportion of girls between 5 and 59%).

Figure 2: Proportion of girls in different specialization tracks



Note: each point in the plot is the proportion of girls in a class of a given cohort and semester.

The relevant outcome variable is student's achievement. Following the previous literature (e.g., Hoxby, 2000; Eisenkopf et al., 2015) we focus our analysis on students' achievements in mathematics and German. Similarly to Eisenkopf et al. (2015), we consider students' achievements as reported by ordinary report cards, since our high school don't run standardized tests at the end of each grade. Table 2 reports average grades in mathematics and German in senior high school (9th to 12th grade) for boys and girls in different specialization tracks exposed to an above/below average proportion of girls in the class.

In the full sample, students in classes with a proportion of girls below the average feature higher average grades in mathematics compared to their fellow students in classes with a proportion of girls above the average (for boys the difference is statistically significant at the 5%-level). This difference, however, partly reflects the selection of boys into the specialization tracks Maths/Science. In fact, classes with a proportion of girls below the average are overrepresented in the specialization track Maths/Sciences. At the same time, in this specialization track, boys tend to have higher grades in mathematics compared to students in other specialization tracks. For grades in German, the tendency is opposite. Boys in classes with a proportion of girls below the average perform worse compared to boys in classes with a proportion of girls above the average. This difference, again, possibly reflects the selection of girls into the specialization track Ancient Languages, where boys tend to have higher grades in German.

When considering each specialization track separately, differences in average grades between students in classes with a proportion of girls above and below the average are also present. Girls in the specialization tracks Modern and Ancient Languages, have higher average grades in mathematics when they are in classes with a proportion of girls above the average. In German, girls tend to have higher average grades when they are in classes with a proportion of girls below the average in all three specialization tracks. On the contrary, boys in all three specialization tracks have higher average grades in German when they are in classes with a proportion of girls above the average. In general, however, differences are small and the null-hypothesis of no difference in means is never rejected.

While table 2 provides first insights into the relationship between gender composition of the class and students' achievements, it fails to take into account important aspects such as students' selection into specialization tracks, differences in teachers' grading and the structure of the data (observations from different semesters are pooled together). Moreover, the binary distinction between classes with a proportion of girls above and below the average might be excessively coarse-grained. Next section discusses a better strategy for the identification of the causal effect of the proportion of girls in the class on students achievements.

Table 2: Average grades in mathematics and German

	Full sample		Modern Languages		Ancient Languages		Maths/ Sciences	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Average grades in mathematics								
Proportion of girls above average	4.11	4.17	4.04	4.17	4.25	4.29	4.25	4.26
	(0.76)	(0.72)	(0.73)	(0.73)	(0.74)	(0.73)	(0.60)	(0.59)
Proportion of girls below average	4.23	4.23	4.01	4.08	4.25	4.19	4.26	4.25
	(0.67)	(0.68)	(0.70)	(0.67)	(0.77)	(0.81)	(0.68)	(0.67)
Difference	-0.12**	-0.06	0.02	0.09	0.00	0.10	-0.01	0.01
Average grades in German								
Proportion of girls above average	4.47	4.60	4.38	4.52	4.61	4.72	4.41	4.53
	(0.46)	(0.40)	(0.42)	(0.40)	(0.58)	(0.41)	(0.44)	(0.44)
Proportion of girls below average	4.41	4.60	4.37	4.55	4.55	4.74	4.38	4.59
	(0.43)	(0.44)	(0.41)	(0.39)	(0.44)	(0.41)	0.43	0.42
Difference	0.06**	0.00	0.01	-0.03	0.06	-0.02	0.03	-0.06
Average share of boys/girls	0.45	0.55	0.28	0.72	0.38	0.62	0.64	0.36

Average grades in mathematics and German in different specialization tracks and in classes with an above/below average proportion of girls in the class. All classes of all cohorts and semesters are pooled.
Standard errors in parentheses.
Grading system: 1 = worst grade / 6 = best grade / 4 = pass.
T-tests for the difference in means: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4 Empirical Strategy

4.1 Identification

There are two major threats to the identification of the causal relationship between gender composition of classes and students' achievements. First, students might select into classes with different gender compositions based on unobservable characteristics that simultaneously affect achievements. This would be the case if, for instance, students (or parents of students) with better analytical skills prefer classes with a lower proportion of girls. Second, teachers might also select into classes with different gender compositions. As an example, if teachers with a stricter grading style tend to select into classes with a higher proportion of girls, the researcher would observe a negative correlation between students' achievements and the proportion of girls in the class, which is merely induced by teachers selection.

Our identification strategy relies on the random assignment of students and teachers to classes within the chosen specialization track in the transition from the 8th to the 9th grade. Random assignment within specialization tracks prevents students and teachers from sorting into classes with a higher or lower proportion of girls based on individual characteristics. Hence, it ensures that the gender composition of the class is uncorrelated with individual characteristics. Similarly to Lavy and Schlosser (2011) and Hoxby (2000), we exploit the variation in the proportion of girls across cohorts within specialization tracks and semesters. Moreover, since we are not concerned with students' sorting into classes, we can additionally exploit variation across classes in the same specialization track and semester. By doing so, we are able to measure the proportion of girls at the class level, which, in our case, constitutes the relevant measure of treatment, since students spend most of the school time with fellow students from the same class.

The base model that we estimate takes the following form:

$$\begin{aligned} grade_{ijts} = &\beta_0 + \beta_1 ratio_{ijts} + \beta_2 female_{ijts} + \beta_3 female_{ijts} * ratio_{ijts} + \beta_4 clsssize_{ijts} \\ &+ \theta_j + \theta_t + \theta_s + \epsilon_{ijts} \end{aligned} \tag{1}$$

where $grade_{ijts}$ is either the grade in mathematics or German of student i, in the specialization track j, with teacher t in semester s. $ratio_{ijts}$ is the proportion of girls in student's i class, $female$ is a gender dummy, $female_{ijts} * ratio_{ijts}$ is an interaction term to test for differential effects on boys and girls, and $classsize_{ijts}$ is the size of student's i class. Because randomization takes place within specialization tracks, we include specialization track fixed effects θ_j. To take into account possible differences in grading among teachers, we include teacher fixed effects θ_t. In doing so, we only exploit that part of variation in the proportion of girls across classes taught by the same teacher within specialization tracks. Finally, we include semester fixed effects θ_s.

Our interest lies in the overall effect as well as in the effect on girls and boys in different specialization tracks separately. In a first step, we estimate (1) in the full sample. We then estimate versions of (1) in different subsamples of girls boys and on students in different specialization tracks separately. Under random assignment of students to classes with different gender composition, the parameter β_1 identifies the causal effect of a higher proportion of girls in the class on grades.

4.2 Validity of the identification strategy

To support the assumption that, conditional on the chosen specialization track, students are randomly assigned to classes, we test for each cohort, whether students assigned to different classes within specialization tracks significantly differ in terms of observable characteristics. To do so, we regress average grades in mathematics and German in junior high school on a battery of 9th grade class fixed effects. If assignment is random, we expect that students in different classes within specialization tracks don't systematically differ in terms of grades in junior high school. Table 3 reports the share of specialization tracks and cohorts, for which the null-hypothesis of jointly non-significance of the class fixed effects could not be rejected at different significance levels.

When we consider the most conservative case (significance level = 0.1), the test detects differences in average grades in junior high school in approximately 19% of the cases for

Table 3: Test for random assignment

Significance level	Grade mathematics	Grade German
0.01	0.92	1
0.05	0.89	0.96
0.1	0.81	0.93

Note: grades in mathematics and German refer to grades during junior high school (7th-8th grades).

mathematics and 7% for German. At a significance level of 0.01, differences in average grades in mathematics in junior high school are found in only 8% of the cases, while classes never differ in terms of German grade. These results support the assumption of random assignment of students to classes.

5 Preliminary results

Table 4 reports the results for the full sample (columns 1-3) as well as for each specialization track separately. In the full sample, we find a positive and significant effect of the proportion of girls in the class on grades in mathematics, but not in German. Specifically, a one standard deviation increase in the proportion of girls in the class is associated with an increase of approximately 0.05 grade points in mathematics. The gender dummy is positive and significant only for the grade in German, indicating that girls' achievements in German are on average higher compared to boys. In mathematics, however, there seems to be no difference between boys and girls.

The results of the regression in each specialization track separately show that the positive effect on grades in mathematics found in columns 1-3 is mainly driven by students in the specialization track Modern Languages (column 4). For German, a higher proportion of girls in the class has a positive impact only on students in the specialization track Ancient Languages. Moreover, the effect is different for boys and girls (i.e., the interaction term is significantly different from zero) in both specialization tracks Ancient Languages and Maths/Sciences.

To dig deeper into the differences between girls and boys in different specialization tracks, table 5 reports the results of the analysis in subsamples. The estimated coefficients reveal that the effects are highly heterogeneous. In the specialization track Modern Languages, the effect on the grade in mathematics is positive for both girls and boys, while the effect on grades in German is not statistically different from zero. In the specialization track Ancient Languages, a higher proportion of girls in the class is associated with higher grades in mathematics and German for girls but not for boys. The effect for girls in the specialization track Maths/Sciences goes in the opposite direction. An increase in the proportion of girls is associated with a decrease in grades in mathematics and German. Moreover, the effect on grades in German is positive for boys.

6 Conclusion

In this paper we investigate the causal relationship between gender composition of the class and students' achievements in mathematics and German using data from one of the largest high schools in the canton of Zurich. In particular, we focus on how the effect varies across students who selected into different specialization tracks.

Previous studies tend to find a positive effect of a higher proportion of girls in the class on achievements of both girls and boys. The policy implications of such findings, however, are limited to distributional considerations. In fact, the gains from increasing the proportion of girls in some classes are, at least partly, offset by losses in the remaining classes.

Our preliminary findings suggest that the effect of a higher proportion of girls in the class is heterogeneous across boys and girls who selected into different specialization tracks.

In particular, while boys and girls in the specialization track Modern Languages (and girls in the specialization track Ancient Languages) profit from having more girls in the class, the effect for girls who selected in the specialization track Maths/Sciences, the effect is negative.

In the discussion on the optimal gender composition of classes, such heterogeneous effects need to be considered. Gathering girls in few classes, or even creating single-sex classes, might only have redistributive effects for some students (e.g., those who selected into the specialization track Modern Languages), since both girls and boys profit from a higher proportion of girls in the class. For more talented girls with possibly, as pointed by (Buser et al., 2014, 2017), a higher taste for competitiveness (e.g., girls who selected into the specialization track Maths/Sciences), this operation might even be harmful. For other students (e.g, those who selected into the specialization track Ancient Languages), this approach might be beneficial, since only girls would profit from it, while boys would be unaffected.

References

Buser, T., Niederle, M., and Oosterbeek, H. (2014). Gender, competitiveness, and career choices. *The Quarterly Journal of Economics*, 129(3):1409–1447.

Buser, T., Peter, N., and Wolter, S. (2017). Gender, competitiveness, and study choices in high school: Evidence from switzerland. *American Economic Review (Papers & Proceedings)*, 107(5):125–30.

Dustmann, C., Ku, H., and Kwak, D. W. (2017). Why are single-sex schools successful? CESifo Working Paper Series 6535, CESifo Group Munich.

Eisenkopf, G., Hessami, Z., Fischbacher, U., and Ursprung, H. W. (2015). Academic performance and single-sex schooling: Evidence from a natural experiment in switzerland. *Journal of Economic Behavior & Organization*, 115:123–143.

Hoxby, C. (2000). Peer effects in the classroom: Learning from gender and race variation. NBER Working Papers 7867, National Bureau of Economic Research, Inc.

Jackson, C. K. (2012). Single-sex schools, student achievement, and course selection: Evidence from rule-based student assignments in trinidad and tobago. *Journal of Public Economics*, 96(1):173–187.

Lavy, V. and Schlosser, A. (2011). Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics*, 3:1–33.

Lazear, E. P. (2001). Educational production. *The Quarterly Journal of Economics*, 116(3):777–803.

Lazear, E. P., Malmendier, U., and Weber, R. (2012). Sorting in experiments with application to social preferences. *American Economic Journal: Applied Economics*, 4(1):136–63.

Oosterbeek, H. and Van Ewijk, R. (2014). Gender peer effects in university: Evidence from a randomized experiment. *Economics of Education Review*, 38(C):51–63.

Park, H., Behrman, J., and Choi, J. (2013). Causal effects of single-sex schools on college entrance exams and college attendance: Random assignment in seoul high schools. *Demography*, 50(2):447–469.

Tables

Table 4: Full sample and students in different specialization tracks

	Full sample			Modern Languages	Ancient Languages	Maths/Sciences
	(1)	(2)	(3)	(4)	(5)	(6)
Mathematics						
Proportion of girls	0.0513**	0.0510**	0.0478**	0.116***	0.0450	−0.0120
	(0.0202)	(0.0204)	(0.0196)	(0.0307)	(0.0389)	(0.0183)
Female	0.0109	0.0109	0.00780	0.00125	−0.0139	0.0214
	(0.0189)	(0.0189)	(0.0189)	(0.0343)	(0.0368)	(0.0284)
Female × proportion of girls	0.0271	0.0271	0.00927	−0.0342	0.0333	−0.0338
	(0.0187)	(0.0188)	(0.0190)	(0.0348)	(0.0436)	(0.0262)
Class size		−0.00473	−0.00948**	−0.0193**	0.00149	−0.0173***
		(0.00411)	(0.00398)	(0.00765)	(0.0102)	(0.00559)
Constant	4.051***	4.185***	4.956***	5.318***	3.508***	4.633***
	(0.0238)	(0.103)	(0.101)	(0.190)	(0.285)	(0.127)
German						
Proportion of girls	−0.00500	−0.00669	0.0212	−0.00668	0.0709***	0.0136
	(0.0153)	(0.0147)	(0.0130)	(0.0214)	(0.0236)	(0.0124)
Female	0.186***	0.187***	0.187***	0.214***	0.185***	0.167***
	(0.0120)	(0.0120)	(0.0119)	(0.0185)	(0.0233)	(0.0206)
Female × proportion of girls	0.000508	0.00261	0.00276	0.0148	−0.0439*	−0.0305*
	(0.0134)	(0.0130)	(0.0124)	(0.0202)	(0.0256)	(0.0178)
Class size		0.00395	0.00188	0.00541	0.00218	0.000658
		(0.00262)	(0.00223)	(0.00451)	(0.00498)	(0.00393)
Constant	4.366***	4.193***	4.032***	4.040***	4.809***	4.104***
	(0.0169)	(0.0630)	(0.0714)	(0.0973)	(0.124)	(0.0932)
Specialization track FE	*yes*	*yes*	*yes*	*no*	*no*	*no*
Semester FE	*no*	*yes*	*yes*	*yes*	*yes*	*yes*
Teacher FE	*no*	*no*	*yes*	*yes*	*yes*	*yes*
Observations	10700	10700	10700	3654	2923	4123

Standard errors in parentheses
Standard errors clustered at the class-semester level
Proportion of girls is standardized with zero mean and unitary standard deviation
Each observation is a student in a semester
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Boys and girls in different specialization tracks

	Modern Languages		Ancient Languages		Maths/Sciences	
	boys	girls	boys	girls	boys	girls
Mathematics						
Proportion of girls	0.121***	0.0816***	−0.00615	0.0849***	0.00565	−0.0879***
	(0.0340)	(0.0248)	(0.0462)	(0.0188)	(0.0164)	(0.0178)
German						
Proportion of girls	−0.00798	0.0136	0.0362	0.0262*	0.0377**	−0.0558***
	(0.0231)	(0.0128)	(0.0307)	(0.0141)	(0.0145)	(0.0170)
Semester FE	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
Teacher FE	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
Observations	1002	2652	1115	1808	2580	1543

Standard errors in parentheses
Standard errors clustered at the class-semester level
Each observation is a student in a semester
Proportion of girls is standardized with zero mean and unitary standard deviation
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$