Erschienen in: Journal of Economic Behavior & Organization. 2015, 115, pp. 123-143
https://doi.org/10.1016/j.jebo.2014.08.004

Academic performance and single-sex schooling: Evidence from a natural experiment in Switzerland☆

Gerald Eisenkopf[a,b], Zohal Hessami[a,*], Urs Fischbacher[a,b], Heinrich W. Ursprung[a]

[a] *University of Konstanz, Germany*
[b] *Thurgau Institute of Economics, Switzerland*

A B S T R A C T

We study the effects of random assignment to coeducational and single-sex classes on the academic performance of female high school students who all face the same curriculum. The students' academic performance is observed over a time period of up to four years. Our estimation results show that single-sex schooling improves the performance of female students in mathematics. This positive effect is particularly large for female students with high ex-ante ability. An accompanying survey reveals that single-sex schooling also strengthens female students' self-confidence and renders the self-assessment of their mathematics skills more level-headed.

Keywords:
Single-sex schooling
Peer effects
Education
Student achievement

1. Introduction

Gender gaps in academic performance, especially in mathematics, continue to be observed worldwide (Guiso et al., 2008; Else-Quest et al., 2010). Since low achievement in mathematics may discourage women from pursuing a career in high-paying occupational fields such as engineering, it is conceivable that the inferior math performance of female students contributes to the persistence of the gender wage gap. The identification of the root causes of gender differences in academic performance is therefore a fundamental economic issue. Especially the relative importance of societal factors as opposed to biological differences influencing the gender gap in mathematics has recently been a focus of economic research.

Our study investigates a particular aspect of the social environment – the gender composition of female students' peer groups in the classroom. So far, the gender composition of peer groups has received little attention in empirical education

☆ The authors thank Lorenz Zubler, Georg Leumann, Claudia Peyer and Constanze Bürki from the Pädagogische Maturitätsschule Kreuzlingen for providing us with anonymous data collected by the school administration. We are also grateful to participants of seminars at the Ruhr Graduate School of Economics, the Thurgau Institute of Economics, the University of Konstanz, the ZEW Mannheim, the Free University of Berlin and the University of Goettingen as well as conference participants in Marbach (Swiss Leading House Economics of Education) and Tübingen (Economics of Education Committee of the German Economic Association). Arye Hillman provided helpful comments. Christoph Kappeler, Martin Reinke and Anne Wunderlich provided excellent research assistance. Finally, we gratefully acknowledge financial support from the Young Scholar Fund of the University of Konstanz.

* Corresponding author at: Department of Economics, University of Konstanz, Box 138, 78457 Konstanz, Germany. Tel.: +49 7531 88 2137; fax: +49 7531 88 3130.

E-mail address: Zohal.Hessami@uni-konstanz.de (Z. Hessami).

economics due to a lack of suitable data. The gender composition often does not vary a great deal across classes or schools, and the data is almost always plagued with (self-)selection problems which make it impossible to identify the causes of the observed differences in academic performance.

One approach taken in the literature is to consider the dichotomy between single-sex education and coeducation which is also the focus of our study. Lee and Lockheed's (1990) study on ninth-grade students in Nigeria, for example, indicates that single-sex schools improve girls' mathematics achievements and engender less stereotype threat in mathematics.[1] The authors acknowledge, however, that a self-selection bias, i.e. differences between the types of students choosing to attend single-sex and coeducational schools, may to some extent be responsible for their result. This is in line with Halpern et al.'s (2011) conclusion that although single-sex schooling "may at first appear promising, apparent advantages dissolve when outcomes are corrected for preexisting differences" (p. 1706).[2] Lee and Lockheed's results also suffer from a second selection bias: in the Nigerian all-girls schools considered in this study, mathematics teachers happen to be exclusively female, implying that the authors cannot isolate gender-specific peer effects from a potential indirect peer effect working through the teacher's gender.

Jackson (2012) exploits data from Trinidad and Tobago where the attendance of single-sex high schools is partially beyond the control of the students. After having taken the Secondary Education Assessment (SEA) exam, the Ministry of Education assigns the students to a high school by using a rule-based mechanism which factors in the students' SEA scores and their preference lists of four schools. Making use of this peculiar institutional setup, Jackson employs a cleverly designed difference-in-difference instrumental variables strategy that isolates the effect of the students' preferred school choice from the effect of single-sex schooling. The results indicate that gaining admission to a preferred school (be it a single-sex or a co-educational school) is associated with better educational outcomes. For most students single-sex education does not appear to provide any additional benefits on top of this positive school-choice effect. Only students expressing very strong preferences for single-sex schooling derive some additional benefits, and among this group girls benefit much more from single-sex schooling than boys. An interesting side effect uncovered by Jackson's study is that girls attending single-sex schools take fewer science courses. One may therefore wonder whether the failure to identify stronger all-girls schooling effects, in particular in mathematics, is a consequence of this course-selection effect working in the opposite direction. It is therefore of special interest to explore the effects of single-sex schooling on the math gender gap in reference groups experiencing identical curricula.

Park et al. (2012) analyze the largest natural experiment on single-sex education so far, the random assignment of South Korean students into single-sex versus coeducational instruction in school districts in Seoul. The authors find that attending single-sex schools, rather than coeducational schools, leads to higher test scores and a higher likelihood of attending four-year colleges. The results hold for boys and girls. However, South Korean coeducational and single-sex schools differ systematically in more than one dimension. Most notably, South Korean single-sex schools are more likely to be private schools with a large degree of autonomy in teacher selection and teacher tenure policies.

While the above studies focus on secondary education, other contributions analyze exogenous variation in gender composition in tertiary education. Booth et al. (2013) analyze pass rates, grades, and course choices in a coeducational university in the UK where students were assigned randomly to coeducational or single-sex classes. The authors observe a positive effect of single-sex education on the performance of female students but no effect on their subsequent course choices. Oosterbeek and van Ewijk (2014) do, however, not find any strong effects at a Dutch university where they exogenously vary the shares of females in workgroups for first-year students in economics and business.

The objective of our study is to follow up the literature on gender differences in educational outcomes by investigating the impact of gender-specific peer effects (single-sex education versus coeducation) on the academic performance of female high school students. Our identification strategy exploits a natural experiment at a high school in the German-speaking part of Switzerland. Just as the vast majority of Swiss high schools, this school is run and financed by the local canton and applies standard curricula and teacher recruitment policies. Since the school has a focus on teaching pedagogics it attracts many female students. In order to provide male students with more peers of their gender in their classes the school board assigns incoming female students to coeducational and single-sex classes. Because of this objective the school board does not apply any specific criteria in this assignment of female students to the different types of classes. Moreover, students and their parents cannot influence the assignment so that the assignment is de facto random. The course program at this school comprises four years, which provides us with a micro panel data-set. All students face at each point of their school career the same curriculum in the core subjects. This renders the students' grades in math and German across parallel classes comparable.[3]

Compared to the traditional research designs of single-sex education studies, we are thus in the fortunate position to perform our investigation in an environment in which the same teachers at the same school teach single-sex and coeducational classes. Moreover, problems of self-selection into classrooms and curricula can be ruled out. Hence, we provide

[1] In this context, stereotype threat represents the experience of anxiety or concern in a situation where a female student faces the risk of confirming the negative stereotype about females' inferior mathematics ability (Steele, 1997).

[2] Halpern et al. (2011) refer to Marsh and Rowe (1996), Thomson and Ungerleider (2004) as well as Smithers and Robinson (2006).

[3] In the first two years, students receive a report card after each semester; in the third and fourth year, they receive only one report card at the end of each year. Hence, up to six grades are recorded for each student per subject.

complementary information to large-scale natural experiments like Park et al. (2012) where single-sex and coeducational schools differ in more than one critical dimension. More specifically, we also learn about the impact of the teachers because we observe the same teachers in different classrooms.

Standard math tests take place in a more competitive environment than, for example, writing essays because grading in math is largely transparent while grading essays depends on tastes that are less transparent. Recent evidence shows that females are equally willing to compete as males if gender stereotypes are not prominent (Shurchkov, 2012).[4] Combined with the common stereotype that math is a "male" subject, this implies that girls in coeducational classes (where the presence of male students makes the stereotype prominent) may suffer from stereotype threat in mathematics but not in native language skills (German). If single-sex schooling indeed reduces or even removes gender-specific stereotype threats, one would expect girls taught in all-girls classes to do better in math than their female peers taught in coeducational classes, but there is no reason to assume that a similar achievement premium will materialize in German.

Our estimation results support this conjecture: we find a positive effect of single-sex education on female students' proficiency in mathematics but not in native language skills (German). In addition, the single-sex class effect in mathematics tends to be stronger for students with high ex-ante ability in math and in classes taught by a male teacher. However, the effect also holds for less talented students and for classes taught by a female teacher. We can show that these effects do not derive from grading-on-a-curve policies. In addition, we do not observe ex-ante differences in math ability between male and female students in our sample which allows us to disentangle the influence of single-sex education on academic performance from a potential ability-driven peer effect.[5]

We additionally conducted a questionnaire survey and show that the identified influence of single-sex education on mathematics achievement is accompanied by changes in the mindset of female students which, in turn, facilitate higher academic achievements: female students educated in single-sex classes, as compared to female students assigned to coeducational classes, evaluate their mathematics skills more positively and are more likely to attribute their performance in mathematics to their own efforts rather than to exogenous talent or luck. Again, in native language skills we do not observe these differences.

The remainder of the study is structured as follows: Section 2 provides a brief survey of related strands of the literature and advances two hypotheses. Section 3 describes the design of the study and the collected data. Section 4 elaborates on the empirical strategy, presents descriptive statistics, and reports the regression results. Section 5 presents the survey-based evidence, and Section 6 offers some conclusions.

2. Related literature and hypotheses

Many potential explanations for the existence of gender differences in academic performance have been explored in the literature. In this section, we briefly discuss the different strands of this literature that are most closely related to our study. We begin with relevant studies on the gender gap in mathematics and then turn to potential sources of the gender gap: gender differences in competitiveness, the role of students' self-perception, and peer-group effects. Based on the insights gained from these studies we advance our hypotheses on gender-specific peer effects.

2.1. The math gender gap

The gender gap in mathematics has recently attracted a great deal of attention in education economics. Guiso et al. (2008) find, for example, that the considerable cross-country variability in the gender gap as measured by the 2003 PISA math test scores is influenced by a socio-economic indicator of gender equality that takes into account females' education opportunities, economic activity, political empowerment, and cultural attitudes toward women.[6] In more gender-equal societies the math gender gap is smaller; the gap even disappears in countries characterized by very high gender equality.

In a large panel data set representative for young schoolchildren in the United States, Fryer and Levitt (2010) find no math gender gap upon entry to school, but substantial differences between boys and girls after six years across every stratum of society. Interestingly, they find little support for the pet hypotheses of many experts maintaining that these differences can be explained by girls investing less effort in the acquisition of math skills, by lower parental expectations, and by biased tests. Fryer and Levitt's study rather confirms the existing cross-country evidence that relates the math gender gap to gender equality at large. These results lead the authors to speculate that the math gender gap is smaller in countries in which schools are gender-segregated, and, as a consequence, they single out this influencing factor as a worthwhile area

[4] Shurchkov (2012) compares the effect of piece rates and tournaments in the context of a mathematical task and a verbal task. He finds that women shy away from competition in the first setting while this cannot be observed in the second setting. He proposes task stereotypes as the explanation for this observed difference.

[5] In a setting where prior to school entry male students have a higher math ability than female students, a positive effect of single-sex schooling on the math performance of female students in single-sex classes might derive from the absence of the (male) students with high math ability. It would not be possible to disentangle an ability effect from a pure gender effect. In our case, male and female students have the same ex-ante ability in math (see Table 3). This allows us to identify a pure gender effect. We thank an anonymous referee for making us aware of this additional advantage of our study.

[6] Cooray and Potrafke (2011) show that the primary determinants of gender inequality in education opportunity are culture and religion, and not political institutions.

for further investigation. Needless to say, cross-country evidence is notoriously difficult to interpret. Furthermore, studies based on evidence gathered from both coeducational and single-sex schools in one country are plagued by serious issues of self-selection. In the light of these considerations, making use of a natural experiment, as we do in our study, may well offer the most convincing identification strategy.

2.2. *Explaining the math gender gap*

2.2.1. *Gender differences in competitiveness*

A large body of literature establishes that men are more willing to compete than women (Gupta et al., 2005; Niederle and Vesterlund, 2007; Gneezy et al., 2009). Sutter and Rützler (2010) even find a gender gap in competitiveness among three-year olds. Gender differences in competitiveness also relate to the math gender gap. Standard math tests take place in a more competitive environment than, for example, writing essays because grading in math is largely transparent while grading essays depends on tastes that are less transparent. More competitive students might therefore prefer to compete on math tests. For this reason, Niederle and Vesterlund (2010) argue that gender-specific attitudes towards competition may cause math test scores to provide a biased picture of true gender differences in math skills, even if the content of these tests is not biased against girls.

Returning to the main focus of our study – the gender composition in the classroom – one may wonder whether females' willingness to compete is affected by their competitor's gender. So far, the evidence on this issue is mixed: Gupta et al. (2013) provide evidence for a significant effect of the competitor's gender in a series of experiments, but Gneezy et al. (2003) and Gneezy and Rustichini (2004) find no effect. Lavy (2013) observes teachers who are placed in a competitive setting in their workplace and also finds no effect of the gender mix on the competitiveness of females. Nevertheless, theoretical arguments support the view that the competitors' gender matters. Steele (1997), for example, introduces the concept of the so-called 'stereotype threat' asserting that females are more likely to conform with gender-specific stereotypes in the presence of males, since they sense gender-specific expectations that they do not want to disappoint.

A potential reason why some of the above studies fail to find a significant influence of the competitor's gender is that such effects may relate to the gender composition of the environment in which the female students are educated over several years. In a field experiment, Booth and Nolen (2012a, 2012b) examine this question. The authors indeed find robust differences between the competitive choices of girls from single-sex and coeducational schools. Moreover, girls from single-sex schools turn out to be more similar in competitiveness to boys even when randomly assigned to mixed-sex experimental groups.

Unfortunately, as pointed out by Niederle and Vesterlund (2010), these conclusions rest on the presumption that the identified behavior of the girls from single-sex schools is not due to the self-selection of more self-assured girls from wealthier families into this type of school. Even though Booth and Nolen go to great lengths to convince the reader that this is not likely to be the case, only a true natural experiment can guarantee that the identified differences in behavior are caused by single-sex schooling.

2.2.2. *Students' self-perception*

Whereas competitiveness plays undoubtedly an important role, other psychological factors may have an even more direct bearing on school achievements. A prime candidate is the *locus of control*, i.e. the way students perceive themselves and their achievements. People with an external locus of control believe that their life is exogenously determined by fate, whereas people with an internal locus of control attribute success and failure to their own actions (Rotter, 1966). Borghans et al. (2008) present experimental evidence showing that individuals with an internal locus of control perform relatively better in cognitive tests. Since the literature suggests that women are more likely to have an external locus of control (Smith et al., 1997), female students may be more easily discouraged from studying hard and acquiring skills. Lee and Bryk (1986) go even one step further and find that this effect depends on the gender of female students' peers. Their study shows that girls in single-sex schools are less likely to blame exogenous factors for their performance.

A second important dimension is the *academic self-concept* which refers to students' self-perceptions regarding their academic achievements (Wigfield and Karpathian, 1991; Ferla et al., 2009). The relationship between academic success and students' academic self-concept and related judgments of self-perceived competence, such as self-confidence, self-esteem, interest, and motivation, is a well-researched issue in educational psychology. Köller et al. (2001) find that students' interest in mathematics at the end of grade 10 has a direct and an indirect effect (via course selection) on achievement in upper-secondary high schools, while other studies (Trautwein et al., 2006a, 2006b) show that ninth-graders' math self-concepts and interests are heavily influenced by the achievements of their peer group, their own achievement, and their grades. Placing students in high-achieving learning groups has, for example, a negative effect on students' academic self-concepts (Trautwein et al., 2009).

Of particular interest for the design of our investigation is the study by Kessels and Hannover (2008) who show in a field experiment that single-sex education in physics improves girls' self-concept of ability. Kessels and Hannover's study does, however, not investigate how single-sex education affects the students' development of cognitive skills. This is the focus of our study.

2.2.3. Peer-group effects

The analysis of peer-group effects has a long tradition in education economics. A number of studies investigate how peer groups defined by race (Link and Mulligan, 1991), parents' education level (McEwan, 2003), social proximity (Foster, 2006), and ability or achievement (Arnott and Rowse, 1987; Lefgren, 2004; Eisenkopf, 2010) affect educational outcomes. Lazear (2001) provides a theoretical analysis that explains why students with behavioral problems end up in smaller classes. Neidell and Waldfogel (2010) find that the unruly behavior of children with limited self-control or discipline has a bad influence on their peers. Since girls and boys often differ in their behavior during secondary education, the gender composition of a class may well have a significant impact on student performance.

Several studies indeed show that the gender ratio in coeducational classes has a strong impact on educational outcomes. Hoxby (2000) using data from Connecticut elementary schools, Whitmore (2005) using American data from the STAR project, and Lavy and Schlosser (2011) using data from Israeli schools observe that an increase in the proportion of girls gives rise to improved student outcomes for boys and girls. These effects are, however, mainly due to having more girls in the classroom and not due to improved behavior of the boys. Black et al. (2013) also find that female Norwegian high school students benefit in the long run from more female classmates, while male students are disadvantaged by female peers. In their critique of scientific studies supporting single-sex education, Halpern et al. (2011) argue that the rather few neurological differences between boys and girls do not suggest different learning styles that would justify a separate education.

The empirical evidence on peer-group effects is, overall, rather mixed and does not lend itself to being easily summarized. In any event, peer-group effects are much harder to identify with rigorous statistical methods than many education professionals appear to assume (see e.g. Manski, 1993, 2000). It is, therefore, all the more remarkable that Schneeweis and Zweimüller (2012) succeed in identifying a causal impact of the gender composition in coeducational classes on female students' choice of secondary school type.

The aggregate evidence leads us to conjecture that the gender composition in the learning environment has immediate effects on the academic performance of female students and circuitous effects working through the differential acquisition of non-cognitive skills. More specifically, we argue that gender differences in attitudes towards specific subjects (e.g. differences in competitiveness in math) induce diverging peer-effects in coeducational and single-sex classes. These differences induce subject-specific stereotype threats and shape the self-perceptions and actual academic performance in that subject. We thus propose:

Hypothesis 1 *(Classmate gender effects in mathematics).* The academic performance of female students in mathematics varies with the gender composition of their classmates.

We deliberately limit our first hypothesis to the academic performance of female students in mathematics since the widely reported math gender gap could have an adverse effect on the self-perception and performance of female students. We conjecture that such a mechanism is not at work in subjects in which female students on average do not perform worse than male students. In these subjects there is, therefore, no generally held prejudice that could give rise to a stereotype threat and female students have a healthy positive self-perception deriving from their well-documented ability to draw level with their male peers if not outperform them. In particular, in native language skills male students perform, as a rule, less well than female students and are therefore less likely to act out their competitive spirit. Moreover, the evaluation of native language skills always entails a good deal of value judgments that do not give rise to a truly competitive situation. Finally, excelling in a subject in which usually girls do well might be frowned upon by a boy's peers as "acting girl-like".

Based on the considerations in the previous paragraphs, one may nevertheless hypothesize that the performance in native language skills will be slightly higher for female students in single-sex classes than for female students in coeducational classes. For instance, this might be due to less classroom disruption. Based on the existing literature which finds no significant effect of single-sex schooling for subjects other than math or physics, we, however, hypothesize that such effects will be very small, if there are any, and we thus propose:

Hypothesis 2 *(Classmate gender effects in native language skills).* The academic performance of female students in native language skills does not vary with the gender composition of their classmates.

To examine this hypothesis we additionally investigate in Sections 3 and 4 the influence of single-sex schooling on skills related to the German language as our study was conducted at a high school in the German-speaking part of Switzerland.[7]

3. Data description

Swiss high schools teach students from 9th to 12th grade and students are between 15 and 19 years old. All high schools in Switzerland offer a fairly standard curriculum. In return, its graduates who pass the school leaving exam (Matura) are automatically allowed to study any subject at any Swiss university. Switzerland, just as Austria and Germany, has a dual education system. After compulsory schooling students either take up an apprenticeship and vocational schooling or they attend a high school. Students typically have to pass a written exam before enrolling at a high school in their respective home canton. In contrast to many other European countries, but also in contrast to Austria and Germany where about one

[7] Pädagogische Maturitätsschule (PMS) Kreuzlingen.

third of a student cohort go to high school, only about 20% of each cohort attend high school in Switzerland. Swiss high schools are therefore still rather elitist institutions for youths who are prepared to make a special effort on their way to a challenging university education.

The high school under investigation makes no exception in this respect. Its four-year curriculum prepares the students to obtaining the school-leaving certificate by covering all the subjects required for the Swiss school-leaving exam. However, it attracts mainly students who, upon graduation, aspire to attend a college of education because the school's curriculum places also emphasis on pedagogical subjects. Having studied these basic pedagogical subjects allows the school's graduates to skip the first-year courses at the University of Teacher Education located in the same town. This arrangement increases the school's attractiveness for students, who intend to become teachers, which, in turn, explains why about 80% of the students are female.

The school board responded to this female-dominated gender composition of the incoming student body by forming girls-only classes in all but one of the eight cohorts that we investigate. According to the school's administration these single-sex classes were introduced in order to increase the share of male students in the mixed classes. Most importantly, the school does not apply any specific criteria to the assignment of incoming students to single-sex and coeducational classes and does not allow for self-selection. When we approached the school board, it was agnostic about the impact of the policy and confirmed repeatedly that the assignment is effectively based on a random process.

This procedure provides a natural experiment on the influence of single-sex education on the academic performance of female students. The experimental setting is quite conservative because the coeducational classes have rather few male students (about 22% rather than the approximately 48%[8] in the other high schools in the canton).[9] We do not claim that students attending this school are perfectly representative for the entire student body in Switzerland, but they are certainly not particularly special. Of course, some students might have chosen the school because its student body is predominantly female. However, we are confident that such a bias would be small (because the main reason for choosing this school is its focus on pedagogical subjects) and will not give rise to a significant change in the relevant results.

Moreover, we do not observe a (self-)selection of students into the school that is based on academic factors. The students do not show an exceptionally strong or poor performance in the standardized exams that determine enrollment into high school education. The school is required to offer a fairly standard curriculum. In return, its graduates, like those from any other Swiss high school, are automatically allowed to study any subject at any Swiss university. We are therefore confident that, at a qualitative level, our results would also hold if female students had been randomly assigned into the school.

We have culled our key data from the school's administrative records which contain information on all 808 students who have joined the school between 2001 and 2008.[10] This includes each student's gender, date of birth, classmates, and report card grades. In each cohort, there are four to five classes with about 18 to 25 students per class. The students assigned to a given class are taught together in all compulsory subjects (about 80% of the total curriculum) for the entire four years of education. Each student takes 12 to 13 courses. Both German (native language skills) and math are compulsory. The first six cohorts include one single-sex class, the seventh cohort includes three single-sex classes and the last cohort has no single-sex class. Upon inquiry, the school board assured us that the variation in the number of single-sex classes is only due to the variation in the share of male students per cohort. The last cohort has the highest ratio of male students (27 out of 106) making it less necessary to introduce a single-sex class to raise the number of male students per coeducational class. On the other hand, cohort 7 has the lowest ratio of male students (13 out of 97) leading to the introduction of three single-sex classes. Table 1 presents the composition of the sample.

Our data allows us to reconstruct across all subjects and semesters by which female or male teacher each student has been taught. Table 2 indicates that single-sex classes were more often taught by female math teachers than coeducational classes. On inquiry, the school management insisted that this outcome certainly does not reflect any intention; it is rather considerations of convenience that underlie the assignment of teachers to classes. In any event, we control in our regression analysis for teacher gender in order to properly identify the peer gender effect. We additionally consider the interaction between the classroom type and teacher gender based on the literature on student–teacher gender interactions and their effects on student behavior and performance (Dee, 2005; Ouazad, 2014).

We also obtained data detailing the incoming students' performance in the entry exams. Students can take the exam at different locations and an overall passing grade allows them to attend any high school in the canton. We obtained the entry exam grades for most students (599 out of 808); earlier cohorts are excluded because of changes in the admission and examination policies. These entry exam grades provide a standardized measure of ex-ante student ability. We use these grades to check whether the assignment to the different class types (single-sex versus coeducational) was actually effected according to a random process as called for by the school's policy statement. The entry exam grades are, of course, also a convenient control measure for ex-ante heterogeneity across female students in single-sex and coeducational classes.

---

[8] We obtained this figure from the statistical office for education in the canton of Thurgau (http://www.bista.tg.ch).

[9] The single-sex literature typically compares two extremes (gender-balanced classrooms vs. single-sex classrooms) while the gender-ratio literature compares slight variations in the gender composition for instance across adjacent cohorts that are quasi-random. Our setting is situated in the middle and is a rather conservative setting that compares all-girls classrooms with coeducation classrooms with just a few male students. This setting makes it likely that we underestimate the effect that one would find in a more traditional coeducation/single-sex environment.

[10] Since these records essentially capture the school board's knowledge about the incoming students, we could actually control for any non-random assignment policy.

Table 1
Distribution of students across cohorts and class types.

	Female students in single-sex classes	Female students in coed classes	Male students in coed classes	Total size of cohort	Number of single-sex classes
Cohort 1 (2001–2005)	19	56	13	88	1 of 5 classes
Cohort 2 (2002–2006)	24	57	15	96	1 of 4 classes
Cohort 3 (2003–2007)	24	71	23	118	1 of 5 classes
Cohort 4 (2004–2008)	18	62	16	96	1 of 5 classes
Cohort 5 (2005–2009)	20	70	18	108	1 of 5 classes
Cohort 6 (2006–2010)	22	62	15	99	1 of 5 classes
Cohort 7 (2007–2011)	52	32	13	97	3 of 5 classes
Cohort 8 (2008–2012)	0	79	27	106	0 of 5 classes
Total	179	489	140	808	9 of 39 classes

Notes: Report cards are handed out twice a year in the first two school years and only once a year at the end of the third and fourth school year.

Table 2
Assignment of female German and math teachers to single-sex and coed classes.

	Math classes	German classes
Single-sex classes	62.0	25.1
Coed classes	31.2	19.6

Notes: Percentages denote the share of students taught by a female teacher (2001–2009).

Table 3 indicates a rather small grade difference in native language skills (German) in favor of female students in coeducational classes but not in mathematics. It is thus not the case that high-ability female students are concentrated in either single-sex or coeducational classes, which would, in any event, be a very unlikely outcome of a random assignment process. Table 3 also indicates that the male students in our sample did not perform significantly better or worse in the qualifying examination than the female students.

The design of our study makes use of the natural experiment deriving from the random assignment of girls to single-sex and coeducational classes. Since the two types of classes have exactly the same curriculum and mode of examination, the random assignment allows a clean identification of how single-sex education of female students influences their academic performance. Tables 8, 9 and 11 in Appendix A describe all of the variables that are included in the empirical analysis.

4. Empirical results

4.1. Descriptive analysis

We measure academic performance with ordinary report card grades because Swiss schools do not run standardized end-of-school-year or exit exams. Grades are, however, a highly incentivized measure as they determine whether a student is promoted to the next grade or retained, and, in the last two school years, grades are an integral part of the matriculation examination. Most importantly, grading is based on criteria that apply to all classes, and the teachers are likely to apply

Table 3
Summary of qualifying exam grades.

	Average grade in math exam (St. dev.)	Average grade in German exam (St. dev.)	Observations
Female students in single-sex classes	3.819 (0.835)	4.124 (0.657)	122
Female students in coed classes	3.824 (0.862)	4.257 (0.667)	375
Difference [*t*-statistic]	−0.005 [−0.049]	−0.133* [−1.925]	497
Male students in coed classes	3.854 (0.879)	4.144 (0.719)	102
Total for all students	3.828 (0.858)	4.210 (0.676)	599

* Significant at the 10 percent level.

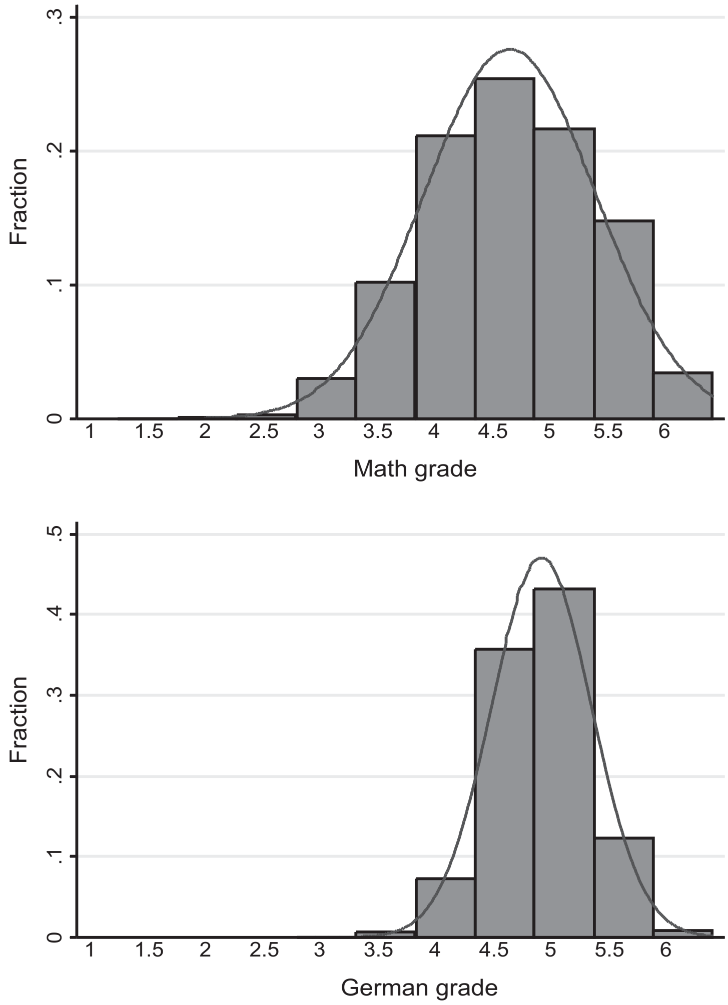

Fig. 1. Distribution of mathematics and German grades of female students.

these criteria very conscientiously since they teach both types of classes.[11] In any event, the application of different standards across the two class types would cause additional costs (e.g. setting different exams or preparing different classes), substantial uneasiness for the teacher, and, given the easy flow of information between students, such a policy would never be sustainable.

Fig. 1 shows the distribution of math and German grades of female students. For each student we observe at most six grades in each subject. In the first two years, students receive a report card after each semester; in the third and fourth year, students receive only one report card at the end of each year. The best grade that can be achieved is 6. Grades decline in steps of 0.5 and 1 is the worst possible grade. Grades of less than 4 in several subjects can lead to retention at the end of a school year.[12] The average grade in mathematics is 4.496 (St. dev.: 0.712) for female students in coeducational classes and 4.665 (St. dev.: 0.738) for female students in single-sex classes. In German classes, the average grade for female students is 4.813 (St. dev.: 0.402) in coeducational classes and 4.807 (St. dev.: 0.431) in single-sex classes.

4.2. Identification strategy

The empirical model is specified as follows. The dependent variable is either the math or the German grade for a specific student on one of the report cards during the four-year period. There are 668 female students that are distributed over 39 classes and that receive up to 6 report cards during the 4 years at this school. All estimations include cohort fixed effects to account for the fact that some cohorts may perform better than others. We successively introduce further explanatory

[11] The nine single-sex classes in our sample were taught by eight math teachers (three female and five male) and twelve German teachers (three female and nine male). Two out of the three female German teachers taught both single-sex and coeducational classes and six out of the nine male German teachers taught both types of classes. In mathematics, all three female teachers taught both types of classes and so did three of the five male math teachers.

[12] Failing grades, i.e. grades below 4, need to be compensated by grades above 4, and compensation is restricted to a limited number of failing grades.

Table 4
Effect of single-sex schooling on female students (mathematics).

Dependent variable: math grade	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Single-sex class	0.162***	0.177***	0.166***	0.172***	0.240***	0.209***
	(3.787)	(3.902)	(3.633)	(3.144)	(3.891)	(3.007)
Age		0.017	−0.070***	−0.070***	−0.070***	−0.070***
		(1.462)	(−3.058)	(−3.082)	(−3.023)	(−3.037)
Class size		0.013	0.007	0.007	0.007	0.008
		(1.414)	(0.801)	(0.793)	(0.765)	(1.010)
Female math teacher				−0.018		
				(−0.266)		
Single-sex class × School year 2						0.076
						(1.643)
Single-sex class × School year 3						0.016
						(0.225)
Single-sex class × School year 4						0.038
						(0.389)
Cohort fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
School year fixed effects	No	No	Yes	Yes	Yes	Yes
Teacher fixed effects	No	No	No	No	Yes	Yes
R^2	0.020	0.018	0.039	0.040	0.040	0.040
Observations	3281	3281	3281	3281	3281	3281
Number of female students	668	668	668	668	668	668

Notes: *t*-statistics are in parentheses. Standard errors are robust to heteroscedasticity and clustered at the class level.
*** Significant at the 1 percent level.

variables such as a student's age, the number of students in the classroom, teacher gender or teacher fixed effects and school-year fixed effects.

To account for the structure of the data, we rely on feasible GLS estimations with random effects at the student-level, and robust standard errors clustered at the class level.[13,14] This procedure allows for a straightforward interpretation of the estimated coefficients. The alternative would be to use ordered probit estimates. Ai and Norton (2003) discuss the interpretation problems related to the interaction effects in logit and probit models estimated with standard statistics programs. They do provide a solution for binary logit and probit models, but not for ordered probit models. We acknowledge that German grades, unlike math grades, are perhaps rather ordinally scaled, depending, of course, on the type of exam and the teacher's grading policy. We therefore re-estimated the regressions presented in Section 4.3 with the ordered probit estimator and briefly discuss the marginal effects of all relevant coefficients except those of the interaction terms.[15]

4.3. *Estimation results*

The estimation results presented in Table 4 suggest that students in all-girls classes obtain better grades in math than their female fellow students in coeducational classes. This effect is rather large and in line with Hypothesis 1.[16] Since virtually all grades range between 3.5 and 6, a coefficient of 0.17–0.24 implies a performance increase of about 7–10% of the relevant range.[17] The cohort fixed effects in model 1 turn out to be jointly significant (the *p*-value corresponds to significance at the 5% level), indicating that there are differences in performance across cohorts.[18] Model 2 additionally includes two control variables: student age and class size. Model 3 also includes school-year fixed effects which are also

[13] Conventional wisdom is that one should cluster at the highest level, i.e. in our case the class level rather than the student level, when those two levels are nested (Cameron et al., 2011). Nevertheless, we have also run the estimations with clustering at the student level and find that the results are, as expected, very similar. Finally, we ran the regressions with clustering at the teacher level; the results are qualitatively the same. For a small number of clusters Monte Carlo evidence suggests that clustering leads more often to erroneous inference than not clustering at all (Cameron et al., 2011), which is the reason why we do not cluster at the teacher-level in the baseline estimations.

[14] Alternatively, we have re-run all of the regressions with the OLS estimator. The differences are quantitatively and qualitatively irrelevant.

[15] Since grades are restricted between one and six, we also ran tobit estimations as a robustness check. The results do not qualitatively differ from those obtained by feasible GLS. Fig. 1 indicates that this result is not surprising given that less than 5% of the students obtained the best grade (6), while nobody received the lowest grade (1).

[16] The estimated effect of 0.17–0.24 translates into about 0.24–0.3 standard deviations. One standard deviation is 0.723 for the female students.

[17] In the corresponding ordered probit estimations of model 4, we find that female students in single-sex math classes have a 1.7% higher probability of obtaining the highest grade of 6 than female students in coeducational classes. This effect is significant at the 5 percent level. In comparison, each year of age reduces the probability of obtaining the highest grade in math by 1.6% (significant at the 1 percent level).

[18] The inclusion of cohort fixed effects also addresses the uneven distribution of single-sex classes across cohorts. As Table 1 shows, especially cohorts 7 and 8 stand out in this respect. By including cohort fixed effects, only the within-cohort variation in performance is considered. Alternatively, we have re-run the regressions in Table 4 excluding the last two cohorts. The results are qualitatively the same (available upon request from the authors). This is not surprising since we have no reason to believe that this variation is due to anything that is correlated with performance.

Table 5
Effect of single-sex schooling on female students (mathematics), subsamples.

Subsample criteria Dependent variable: math grade	Female math teacher Model 7	Male math teacher Model 8	Full sample Model 9	Student in single-sex class Model 10	Student in coed class Model 11
Single-sex class	0.134***	0.306***	0.323***		
	(3.054)	(3.871)	(3.887)		
Age	−0.118***	−0.050*	−0.070***	−0.083*	−0.066**
	(−3.207)	(−1.823)	(−3.066)	(−1.920)	(−2.455)
Class size	0.013	0.015	0.008	0.022*	0.001
	(0.691)	(1.302)	(0.863)	(1.870)	(0.052)
Female math teacher			0.033	−0.396***	0.039
			(0.552)	(−5.024)	(0.626)
Single-sex class × Female math teacher			−0.276***		
			(−2.821)		
Cohort fixed effects	Yes	Yes	Yes	Yes	Yes
School year fixed effects	Yes	Yes	Yes	Yes	Yes
R^2	0.056	0.041	0.038	0.041	0.033
Observations	1316	1965	3281	898	2383
Number of female students	366	444	668	180	489

Notes: *t*-statistics are in parentheses. Standard errors are robust to heteroscedasticity and clustered at the class level. In models 7 and 8, the sum of female students is larger than 668 since some students were taught by both male and female teachers. In addition, with regard to models 10 and 11 there is one student who switched classes.
* Significant at the 10 percent level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level.

jointly significant. Model 4 additionally includes a dummy for female math teachers and reveals that the teacher's gender is not driving the teacher-specific grading differences. Further robustness checks (see Tables 15 and 16 in Appendix B) also show that the grading differences are not driven by idiosyncratic characteristics of a specific teacher or a particular single-sex class and that the effect is even slightly larger (single-sex class coefficient between 0.22 and 0.25) when only considering the performance assessed by math teachers that teach both single-sex and coeducational classes within the same cohort and in the same year/semester (results available upon request). Furthermore, in model 5, we include teacher fixed effects instead of a teacher gender dummy, i.e. we test whether female students in single-sex classes perform better than their counterparts in coeducational classes even if both classrooms have the same teacher.

In model 6, we include interaction terms between class type and school years. We find no evidence that the positive single-sex class effect on female students' performance steadily increases as the students advance to higher grades. However, we observe that in the second year some additional benefits of single-sex education materialize even though they fall short of conventional significance levels. It thus appears to take some time for the beneficial effects of single-sex schooling to consolidate. We discuss this issue in greater detail in Section 4.4.3 where we focus the analysis on the freshman students. The result is also perfectly in line with our survey-based evidence reported in Section 5: single-sex education has a decided effect on math-related psychological traits, but these effects also do not form immediately.

The coefficient estimates of the remaining control variables included in the regressions illustrate that older students of a given cohort in a given school year perform worse. The age effect can arguably be attributed to academically weaker students having lost a year or two before entering the school (late enrollment to primary school, academic detours, etc.). Class size does not appear to have an influence on academic achievements.[19]

As already pointed out, model 4 shows no direct teacher gender effect. In Table 5, we investigate the teacher gender effect more closely: we examine how teacher gender affects the impact of class type on student performance. Models 7 and 8 reveal that single-sex schooling benefits female students regardless of teacher gender. However, the effect is smaller for female teachers. In model 9, we report results for a regression including an interaction term of class type and teacher gender. The significance of this term shows that male and female teachers have indeed a different impact on the peer-gender effect. In models 10 and 11, we report the teacher gender effect separately for single-sex classes and coeducational classes. These models show that the students' academic performance is only (negatively) affected by female teachers in all-girls classes. In summary, even though teacher gender has no influence on grading per se, there is an interaction of teacher gender and the gender-specific peer effect – teacher gender affects the academic achievements of students in all-girls classes.

We also tested whether the academic performance of male or female students improves if the number of male students in a coeducational class gradually increases. Yet, we do not observe such an effect.[20] We are therefore led to conclude that it

[19] Studies analyzing the degree to which class size matters for student achievement have provided rather mixed results. See Rockoff (2009) for a survey of early 20th century field experiments and a summary of the more recent literature.

[20] We have coeducational classes with 2 to 8 male students. We have conducted an additional estimation (not shown here) where we include dummies for the number of boys in the class (base category is 0 boys). We find a significantly negative effect for all of the dummy coefficients, while the sizes of the coefficients are of the same order of magnitude.

Table 6
Effect of single-sex schooling on female students (German).

Dependent variable: German grade	Model 12	Model 13	Model 14	Model 15	Model 16	Model 17	Model 18
Single-sex class	0.005	−0.001	−0.002	−0.002	−0.008	0.001	−0.012
	(0.236)	(−0.042)	(−0.124)	(−0.111)	(−0.283)	(0.016)	(−0.415)
Age		0.015**	−0.039***	−0.039***	−0.039***	−0.039***	−0.039***
		(2.039)	(−3.003)	(−2.985)	(−3.132)	(−3.132)	(−2.964)
Class size		−0.006	−0.005	−0.005	−0.007	−0.007	−0.006
		(−0.763)	(−0.623)	(−0.601)	(−0.923)	(−0.867)	(−0.683)
Female German teacher				−0.001			−0.016
				(−0.031)			(−0.266)
Single-sex class × School year 2						−0.010	
						(−0.192)	
Single-sex class × School year 3						0.026	
						(0.507)	
Single-sex class × School year 4						−0.068	
						(−1.114)	
Single-sex class × Female German teacher							0.036
							(0.519)
Cohort fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School year fixed effects	No	No	Yes	Yes	Yes	Yes	Yes
Teacher fixed effects	No	No	No	No	Yes	Yes	No
R^2	0.006	0.005	0.034	0.034	0.046	0.047	0.034
Observations	3281	3281	3281	3281	3281	3281	3281
Number of female students	668	668	668	668	668	668	668

Notes: *t*-statistics are in parentheses. Standard errors are robust to heteroscedasticity and clustered at the class level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level.

is the very absence of male students that drives our results. At a first glance, this result appears to be at variance with a recent finding by Lavy and Schlosser (2011) who suggest that an increase in the proportion of girls improves cognitive outcomes of both boys and girls. Since the effect identified by Lavy and Schlosser works through less classroom disruption when the share of girls is high, this channel of influence is not likely to be relevant in our elitist Swiss high school environment which, moreover, is dominated by students who aspire to become teachers themselves.

We now turn to the analysis of the academic performance in native language skills (German). Table 6 reports our estimation results. The estimates indicate that in German students in all-girls classes do not outperform students instructed in mixed classes. Moreover, teacher gender has no impact on female students' performance, neither in single-sex nor in coeducational classes. We observe negligible benefits of a reduction of class size which is in line with the evidence by Wößmann and West (2006) for countries with relatively high teacher salaries.[21]

4.4. Additional specifications

4.4.1. Ex-ante ability and effect heterogeneity

Models 19 to 21 in Table 12 (see Appendix B) replicate models 3 and 4 in Table 4 and model 9 in Table 5 to check whether controlling for ex-ante abilities as measured by the grades received in the qualifying exam has an influence on our estimates. This is not the case. We have not included this variable in the baseline estimations as we only have data for this variable for 497 female students.[22] The grades from the qualifying exam have, not surprisingly, a strong explanatory power for the students' subsequent academic performance and capture a substantial share of the ex-ante heterogeneity among the student body.

To investigate potential heterogeneity across ability levels, we have estimated models in which we interact ex-ante ability (as measured by the math grade in the qualifying exam) with our single-sex class dummy. In model 22, we carry out this exercise for the full sample and in models 23 and 24 we distinguish between female students that were taught by a male or female math teacher. The interaction term in model 22 turns out to be positive and significant at the 10 percent level. Hence, female students with high ex-ante math abilities benefit more from single-sex education in math than academically weaker students. Compare, for example, two female students who are identical in all respects with the exception of one having performed better by one grade point in the math qualifying exam. If both attend single-sex classes, the more able one performs better in math by about 45% of a grade point than the less able one. If the two girls had both been assigned to

[21] In the corresponding ordered probit estimations with the German grade as the dependent variable, we also find that the coefficient and the marginal effect of the single-sex class dummy is insignificant.

[22] We also included the qualifying exam grade in German as an additional control variable in the regressions shown in Table 6. Also in this case the baseline results are not affected.

a coeducational class, the difference would have been only 31% of a grade point. The estimated overall effect of single-sex schooling is in line with the results of our prior estimations: for an average grade in the qualifying exam for female students in math, i.e. about 3.8 (see Table 3), the improvement in the math grades due to single-sex schooling amounts to 0.16 based on the results in model 22.[23]

The result that good math students benefit more from single-sex education supports the view that relates the beneficial effects of all-girls schooling to the absence of gender-specific stereotype threats. If the stereotype suggests that high performance in math does not agree with womanly qualities, female students with a potential for above average achievements in math are especially impaired. As soon as the stereotype threat disappears, these female students will begin to thrive, whereas below average students, who were never in danger of violating the stereotype, have no reason to change their behavior.

4.4.2. Grading on a curve

Even though it would counteract explicit school policy, it is not entirely inconceivable that some teachers grade according to a predefined distribution that is imposed on each class. If male students performed better than female students, female students in single-sex classes would, under a grading-on-a-curve policy, obtain on average better grades than in a coeducational class as there are no male students present to capture the highest grades.

In Table 13 in Appendix B, we report the results of four regressions based on the grades of male *and* female students. Models 25 and 26 provide evidence that single-sex classes perform better in math than co-educational classes even if male students are taken into account. This can be inferred from the single-sex coefficient which is significant at the 1 percent level in both cases. Models 27 and 28 do not indicate any differences between single-sex and coeducational classes in terms of native language skills. This result clearly shows that the single-sex education effects that we have identified are not an artifact of grading on a curve.[24]

4.4.3. Treatment effects in the first year

Another concern might be that our results are influenced by dropouts and grade retention. Since we argue that single-sex education improves the girls' academic performance, it is less likely that students assigned to all-girls classes have to repeat a year or are even forced to drop out. The random assignment could therefore become somewhat diluted after the first school year, i.e. after the first two semesters. We have therefore re-estimated our baseline models using only the grades earned in the first two semesters. The results are summarized in Table 14 in Appendix B. The sample size is slightly smaller than in the baseline estimations since some students join this school only in the second year.

When restricting the sample to the first year of studies (when weak students have not yet dropped out or repeated a year), we continue to find a significant single-sex effect. The coefficient is however somewhat smaller than in the baseline estimates. This is presumably so because it takes time for the full benefits of single-sex schooling to materialize (see our comments on model 6 and the survey evidence reported in Section 5). Another finding is that the interaction between the single-sex class dummy and the female math teacher dummy is small in comparison to model 9 in Table 5 and not significant at this early stage. Since all math teachers but one taught at this level the teacher assignment procedure does not explain this difference.

4.4.4. Summary of the additional results

Overall, the results from the additional specifications support our first hypothesis: the mathematics performance of female students varies with the gender composition of their classmates. Three qualifications with respect to these direct gender-specific peer effects are however called for. First, the relationship between the gender composition of the class and the academic performance of female students appears to be highly non-linear: the mere presence of male students compromises the educational environment that is especially conducive to the female students' academic development. Second, the effect of single-sex education on the math performance of female students is positive across all performance levels, but it is larger for good math students. Third, we find that the teacher gender has an impact on the size of the treatment effect. In all-girls classes, all teachers are able to elicit better accomplishments in mathematics but male teachers induce a stronger effect.[25] Since we do not observe this difference in the first year, the result suggests that the positive peer effect in single-sex classes shapes the behavior of male teachers who increase their own teaching efforts accordingly and accentuate the peer effect even more. Moreover, single-sex education is not advantageous to female students across the board. The results also support our second hypothesis. The performance of female students in native language skills does not vary with the gender composition of their classmates.

---

[23] The estimates of model 20 also imply that the effect of single-sex schooling on math grades is almost always positive: the negative single-sex class effect is at least compensated by the positive interaction term as long as the qualifying exam grade of a student is 2.7 or higher ($-0.368 + 0.138 \times 2.7 > 0$).

[24] We also estimated model 19 at the class-level. The results are very similar.

[25] This result raises two questions. First, why do female high-school students respond positively to male teachers, whereas female college students derive benefits from female instructors as documented by Carrell et al. (2010)? Is it the age of the students or the different classroom environment (single-sex versus coeducational)? Second, do male students in a single-sex education environment also respond to the gender of their teachers? In our sample we can, of course, not investigate these questions.

Table 7a
Survey responses by female students (10th to 12th grade, i.e. Cohorts 6 to 8) attending single-sex and coeducational classes in spring 2010.

	Math self-concept		Math self-assessment		German self-assessment	
	Observations	Response	Observations	Response	Observations	Response
Coed	147	3.051	150	2.032	145	2.785
Single-sex	61	3.402	61	2.382	60	2.850
Difference		−0.351		−0.350*		−0.065
(*t*-statistic)		(−1.458)		(−1.828)		(−0.466)
Total	208	3.154	211	2.133	205	2.804

* Significant at the 10 percent level.

Table 7b
Survey responses by female students (9th grade, i.e. Cohort 9)[a] attending single-sex and coeducational classes in spring 2010.

	Math self-concept		Math self-assessment		German self-assessment	
	Observations	Response	Observations	Response	Observations	Response
Coed	64	3.254	64	2.199	65	2.762
Single-sex	20	3.238	19	2.184	19	2.842
Difference		0.016		0.015		−0.080
(*t*-statistic)		(0.044)		(0.045)		(−0.317)
Total	84	3.250	83	2.196	84	2.780

[a] Cohort 9 was not included in the preceding regression analysis since we do not have any administrative data for these students (e.g. grades, age, teachers, etc.).

5. Survey evidence

Studies in social psychology typically reveal a strong relationship, usually interpreted to be mutually reinforcing, between subject-specific ability and related assessments of self-perceived competence (Köller et al., 2001; Trautwein et al., 2006a, 2006b, 2009). Because of this established link between self-perceived competence and performance, we conducted a survey among the currently enrolled students in order to check whether single-sex schooling actually influences the students' self-assessment and thereby, presumably, academic achievement. If such positive effects are observable, this might suggest that single-sex schooling has an effect on female students' mindset and their academic performance after high school.

This survey was conducted in March 2010 and covers 213 female students, 62 of which were enrolled in single-sex classes and 151 in coeducational classes. The teachers administrated the survey, and the students answered the questions in an ordinary lesson without receiving any information whatsoever on the purpose of this survey. The survey comprised questions about students' family background and their attitudes towards mathematics and German.

Our survey also shows that there is no worrying statistical relationship between the students' socio-economic family background and their assignment to the two types of classes. Out of 25 family background characteristics only three indicated a significant correlation (at the 10% level) with the girls' assignment to the two types of classes (see Table 10A in Appendix A): the families of girls assigned to single-sex classes own fewer musical instruments and these girls' parents are more likely to have had vocational training (as compared to no training or higher education). If anything, this might hint at a slightly lower economic status of the families of girls assigned to single sex-classes.[26] If this were indeed the case, our results would actually be even stronger since it is well known that a low socio-economic status has a detrimental effect on academic achievement (Schütz et al., 2008).

To operationalize the different concepts portraying the students' mindset, we employed psychological scales that have been widely used in educational psychology. Students were asked to divulge how much they agreed with nine different statements (on a scale from 1 to 4). Five statements captured the math-specific self-concept (which measures the perceived relationship between effort of studying and success) and the remaining four statements measured the self-assessment of the student's own skills in math (see Table 9 in Appendix A).[27] Not surprisingly, both measures turn out to be highly correlated with performance.

Tables 7A and 7B summarize the main results. Among the female students, we observe a stronger self-concept in mathematics and a more positive self-assessment of mathematics skills in single sex-classes than in coeducational classes. In

[26] We have also computed the entries of a table similar to Table 10A which analyzes whether the assignment of female and male teachers to single-sex classes is influenced by the students' socioeconomic background (see Table 10B in Appendix A). We find for only two of the 25 variables (number of computers and number of musical instruments at home) a statistically significant difference at the 10 percent level. This result is also perfectly in line with random teacher-student matching.

[27] We use the same statements that have been used in relevant psychological studies in German speaking countries (e.g. Köller et al., 2000, 2001). The relevant statistical procedures and measures (principal component analysis or Cronbach's alpha) provide empirical support for the integration of these statements into two scales.

German, we do not observe any differences in self-assessment across the two class types. Nor is there any difference among the first-year students, indicating that getting rid of long-held views and attitudes takes time.[28]

We re-estimated model 1 in Table 4 by restricting the sample to the current student population and included the math-related psychological measures as control variables. The effect of single-sex education on performance remains significant.

We acknowledge, of course, that we cannot cleanly identify the causal relationship between these mathematics-related psychological traits and math performance. Our observations concerning self-concept and self-assessment are, however, compatible with the existence of a channel of influence running from the educational environment to the student's mindset which, in turn, affects her academic performance. This circuitous channel of influence does, of course, not exclude a more direct effect of single-sex education on academic performance. As a matter of fact, our empirical evidence is suggestive of such a direct effect which is, moreover, likely to amplify the psychological effect because better performance helps to build up self-confidence.

In any event, we conclude that the described influence of single-sex education on the female students' mindset is an important by-product of the identified correlation between single-sex education and academic performance because this mechanism is in line with the accumulating evidence that single-sex education engenders a specific kind of social learning. Single-sex education appears, for example, to give rise to more competitive behavior (Booth and Nolen, 2012a) and lower levels of risk aversion (Booth and Nolen, 2012b). Note, finally, that an enhanced self-confidence of students educated in single-sex classes can be beneficial in itself since it renders female students less reluctant to choose further education in challenging subjects (see, for example, Compte and Postlewaite, 2004; Schneeweis and Zweimüller, 2012).

6. Conclusion

Based on insights gained from pedagogical considerations, many educators have arrived at the conclusion that single-sex education in "male" subjects such as mathematics and science may be advantageous for female students. Unfortunately, only little experience with single-sex education has been gathered in the more recent past, and, more importantly, the information deriving from these experiences cannot easily be converted into meaningful investigations because comparisons across school types are fraught with the suspicion of being contaminated with self-selection problems: it is virtually impossible to rule out that girls or their parents who opt voluntarily for an all-girls school are not special in some unobservable characteristics. Up to now, convincing empirical evidence concerning the effects of single-sex education has therefore been rather limited and focused on differences across schools which differ in many more dimensions.

We provide the first evaluation of female single-sex education with a randomized assignment of female students from the same school and with the same teachers into different learning environments and find strong empirical support for the benefits of single-sex education. Analyzing a natural experiment performed at a high school in Switzerland, we estimate the impact of single-sex education on the academic performances of female students. We find a positive effect of single-sex education on the proficiency in mathematics but not in native language skills. The effect in mathematics is robust for all observable circumstances. It tends to be stronger if girls in a single-sex class are taught by a male teacher and if girls have a high ex-ante ability in math.

Since our results are derived from a natural experiment, they are not likely to be subject to any selection bias. In order to support our claim that selection does not play any role in our findings, we apply two robustness checks. First, we show that the female students attending single-sex classes in our sample school are not different from the ones attending mixed classes. Moreover, the homogeneity presumption with respect to the student body across the two control groups is also supported by the fact that controlling for ability or initial academic knowledge as measured by a standardized entry test does not change our results. Second, we show that single-sex instruction in mathematics outperforms instruction in mixed classes even if the performance of all (male and female) students attending mixed classes is used as the basis of comparison. This rules out that our findings are an artifact of an implicit grading-on-a-curve policy, or of gender-based discrimination.

It remains to discuss the likely causes for the empirically identified single-sex schooling effect. The fact that the effect only materializes in mathematics but not in German may hint at the underlying mechanisms. We propose two hypotheses. The first one is not novel and derives from the simple observation that girls may suffer from stereotype threat in mathematics but not in German. If single-sex schooling indeed reduces or even removes gender-specific stereotype threats, one would expect girls taught in all-girls classes to do better in math than their female peers taught in coeducational classes, but there is no reason to assume that a similar achievement premium will materialize in German since this subject is not fraught with such a threat. Our results based on data from a questionnaire survey are also in line with this interpretation.

Whether the stereotype threat paradigm explains the identified teacher–classroom type gender interaction effect, is however questionable. To be sure, it is conceivable that the math-anxiety of (female) teachers may carry negative consequences for the math achievement of their female students. But this effect has only been observed for primary school teachers (Beilock et al., 2010); it is not likely that female high-school teachers who have studied mathematics at the university level are afflicted with this kind of anxiety. Dee (2005) finds that performance and behavior evaluations of 8th-grade students by their teachers are more negative when the teacher is of the opposite sex, while Ouazad (2014) does not find significant

[28] There was no single-sex class for those students who were in the second year of their studies at the PMS in spring 2010.

student–teacher gender interaction effects for students in elementary school. The first study mentions as a potential explanation for the significant student–teacher gender interaction that the stereotype threat might be more prominent when female students are taught by male teachers in "male" subjects. In the context of our setting, this would imply a weaker effect of single-sex schooling on math performance when an all-girls class is taught by a male teacher rather than a female teacher. This is, however, the opposite of what we find.

Our second hypothesis is based on a speculation that derives from the notion that peer-competition is a major driving force behind the effort exerted by high-school students. Mathematics, as it is taught at high schools, is a subject that allows applying objective and cardinal performance measures – and such measures are also routinely applied. This kind of grading is not open to ifs and buts and therefore invites outright competition. High school instruction in the mother tongue – in our case German – focuses, on the other hand, on writing essays, reading and discussing a certain canon of literature. These are many-dimensional skills and even the individual dimensions are hardly amenable to objective evaluation. Evaluations therefore stand on shaky ground. As a consequence, evaluations often turn out to be rather non-committal and reflect a great deal of caution. The grade distributions reported in Fig. 1 support this contention. Since mathematics instruction invites competitive behavior, it is not surprising that pubescent boys welcome this opportunity to touting their prowess – especially in the company of girls. Girls that are on average less competitive, on the other hand, are likely to refrain from trying too hard because they know that the boys are committed to high effort. As soon as boys, i.e. contestants committed to high effort, are not present any more, competition becomes more rewarding for the girls. The girls are therefore likely to spend more effort in single-sex classes and accordingly perform better.[29]

Even though the identified positive effect of single-sex schooling in mathematics appears to be very robust, the consequences for education policy remain unclear. Before drawing far-reaching conclusions we need to better understand the mechanisms underlying the identified effect. Moreover, we identify three issues as important avenues for future research. First, it remains to be seen whether female students who perform better in math due to single-sex schooling are likely to choose different subjects in college or different careers. Second, it is of interest to see how these females perform in a potentially male-dominated work environment when they grow older. Third, any policy that propagates all-girls classrooms also has an indirect influence on the male students' study environment. It would therefore be interesting to learn more about the impact on male students. For our rather small sample of male students, we find preliminary evidence that the share of the female students in a coeducational class (which varies between 65% and 88% in our sample) has no significant impact on the performance of male students (results available upon request).

Appendix A. Data description

See Tables 8–11.

Table 8
Definitions of variables.

Variable	Description
Dependent variables	
Math grade	Respective grade in each report card measured
German grade	on a scale running from 1 (very bad) to 6 (very good) with 0.5 steps
Student-level control variables	
Female student	Gender dummy for students (1: Female, 0: Male)
Age	Age of student in full years when report card was handed out
Cohort	Dummies for the student cohorts
School year	School year in which report card was handed out
Math grade in qualifying exam	Grade for standardized written examination in mathematics
German grade in qualifying exam	Grade for standardized written examination in German
Class-level control variables	
Class size	Total number of students in a class
Single-sex class	Dummy for type of class (1: All-girls, 0: Coeducational)
Math teacher	Dummies for the mathematics teachers
Female math teacher	Gender dummy for mathematics teachers (1: Female, 0: Male)
German teacher	Dummies for the German teachers
Female German teacher	Gender dummy for German teachers (1: Female, 0: Male)

[29] See Oosterbeek and van Ewijk (2014) for survey data evidence on peer effects on competitiveness in the context of their study in which they vary the gender ratio of student workgroups in business and economics at a Dutch university.

Table 9
Survey questions on students' self-perception.

Variable	Items
Math self-concept	1. I would enjoy doing math more if it were not so difficult. 2. Even though I try hard, it appears more difficult for me than for my fellow students to study math. 3. Nobody is good at everything. I simply have no talent for math. 4. With regard to some questions in math that I did not understand, I know right away: "I will never understand this." 5. I do not have a particular talent for math.
Math/German self-assessment	1. I often worry that the math/German classes are too difficult for me. 2. I am just not good at math/German. 3. I find it easy to study math/German. 4. In my math/German classes I even understand the most difficult questions.

Notes: Items are rated on a four-point Likert scale ranging from (1) "Strongly agree" to (4) "Strongly disagree". We used the following headline question: To what extent do you agree with the following statements? Math self-concept is scaled from 0 to 5 and is based on the five items in the upper panel of this table. Each item yielded 0.25, 0.5, 0.75 or 1 point (weaker agreement with the statement yielding a higher score). The same method was applied for the 0 to 4 scaling of the self-assessment variables. Cronbach's alpha amounts to 0.919, 0.907, and 0.855 for the math self-concept, math self-assessment and German self-assessment, respectively.

Table 10A
Mean values for socioeconomic background, female students by class type.

Class type	Born abroad	Obs	Mother born abroad	Obs	Father born abroad	Obs	German is native language	Obs	Parents separated/divorced	Obs
Coed class	0.047	150	0.107	150	0.140	150	0.974	151	0.175	149
Single-sex class	0.048	62	0.177	62	0.145	62	0.952	62	0.129	62
Difference	−0.001		−0.070		−0.005		0.022		0.046	
(*t*-statistic)	(−0.054)		(−1.406)		(−0.098)		(0.812)		(0.816)	
Total	0.047	212	0.127	212	0.142	212	0.967	213	0.161	211

Class type	Number of books	Obs	Number of musical instruments	Obs	Musical instrument (Y/N)	Obs	Number of rooms	Obs	Number of bathrooms	Obs
Coed class	3.715	151	6.093	151	0.987	151	7.635	137	2.139	151
Single-sex class	3.548	62	4.952	62	1.000	62	7.276	58	1.968	62
Difference	0.167		1.141*		−0.013		0.359		0.171	
(*t*-statistic)	(1.109)		(1.884)		(−0.908)		(0.910)		(1.039)	
Total	3.667	213	5.761	213	0.991	213	7.528	195	2.089	213

Class type	Number of cars	Obs	Number of computers	Obs	Number of TVs	Obs	Number of cell phones	Obs	Number of siblings	Obs
Coed class	1.818	148	3.450	151	1.551	147	4.490	151	2.080	151
Single-sex class	1.936	62	3.532	62	1.433	60	4.419	62	2.113	62
Difference	−0.118		−0.082		0.118		0.071		−0.033	
(*t*-statistic)	(−0.687)		(−0.342)		(0.800)		(0.340)		(−0.165)	
Total	1.852	210	3.474	213	1.517	207	4.470	213	2.089	213

Class type	Mother has no degree	Obs	Mother had vocational training	Obs	Mother has secondary certificate	Obs	Mother has high school degree	Obs	Mother has a college degree	Obs
Coed class	0.020	151	0.576	151	0.020	151	0.219	151	0.166	151
Single-sex class	0.000	61	0.721	61	0.033	61	0.131	61	0.115	61
Difference	0.020		−0.145**		−0.013		0.088		0.051	
(*t*-statistic)	(1.107)		(−1.978)		(−0.559)		(1.459)		(0.933)	
Total	0.014	212	0.618	212	0.024	212	0.193	212	0.151	212

Class type	Father has no degree	Obs	Father had vocational training	Obs	Father has secondary certificate	Obs	Father has high school degree	Obs	Father has a college degree	Obs
Coed class	0.007	149	0.584	149	0.000	149	0.121	149	0.289	149
Single-sex class	0.000	61	0.721	61	0.016	61	0.066	61	0.197	61
Difference	0.007		−0.137*		−0.016		0.055		0.092	
(*t*-statistic)	(0.639)		(−1.873)		(−1.568)		(1.185)		(1.374)	
Total	0.005	210	0.624	210	0.005	210	0.105	210	0.262	210

[1] The values for the variable 'number of books' have categorical meaning ranging from 1 (0–10 books), 2 (11–25 books), 3 (26–100 books), 4 (101–200 books), 5 (201–500 books) to 6 (more than 500 books) in the household. [2] The variables concerned with the parents' educational background refer to the highest degree obtained by each parent. [3] We also analyzed whether there are significant differences with regard to the number of younger brothers/sisters, equally old brothers/sisters and older brothers/sisters. The *t*-statistics are insignificant in all six cases. [4] *t*-statistics are in parentheses.

* Significant at the 10 percent level.
** Significant at the 5 percent level.

Table 10B
Mean values for socioeconomic background, female students by gender of math teacher.

Class type	Born abroad	Obs	Mother born abroad	Obs	Father born abroad	Obs	German is native language	Obs	Parents separated/divorced	Obs
Male math teacher	0.034	87	0.128	86	0.105	86	0.954	87	0.184	87
Female math teacher	0.065	108	0.138	109	0.174	109	0.972	109	0.138	109
Difference	−0.031		−0.010		−0.069		−0.018		0.046	
(*t*-statistic)	(−0.952)		(−0.197)		(−1.377)		(−0.689)		(0.880)	
Total	0.051	195	0.133	195	0.144	195	0.964	196	0.158	196

Class type	Number of books	Obs	Number of musical instruments	Obs	Musical instrument (Y/N)	Obs	Number of rooms	Obs	Number of bathrooms	Obs
Male math teacher	3.770	87	6.483	87	0.989	87	7.720	75	2.069	87
Female math teacher	3.633	109	5.367	109	0.991	109	7.417	103	2.138	109
Difference	0.137		1.116*		−0.002		0.303		−0.069	
(*t*-statistic)	(0.957)		(1.896)		(−0.160)		(0.775)		(−0.424)	
Total	3.694	196	5.862	196	0.990	196	7.545	178	2.107	196

Class type	Number of cars	Obs	Number of computers	Obs	Number of TVs	Obs	Number of cell phones	Obs	Number of siblings	Obs
Male math teacher	1.753	85	3.207	87	1.536	84	4.356	87	2.172	87
Female math teacher	1.889	108	3.624	109	1.472	106	4.587	109	2.092	109
Difference	−0.136		−0.417*		0.064		−0.231		0.081	
(*t*-statistic)	(−0.839)		(−1.836)		(0.462)		(−1.165)		(0.409)	
Total	1.829	193	3.439	196	0.069	190	4.485	196	2.128	196

Class type	Mother has no degree	Obs	Mother had vocational training	Obs	Mother has secondary certificate	Obs	Mother has high school degree	Obs	Mother has a college degree	Obs
Male math teacher	0.011	87	0.621	87	0.023	87	0.218	87	0.126	87
Female math teacher	0.019	108	0.648	108	0.028	108	0.157	108	0.148	108
Difference	−0.008		−0.027		−0.005		0.061		−0.022	
(*t*-statistic)	(−0.394)		(−0.394)		(−0.209)		(1.089)		(−0.434)	
Total	0.015	195	0.636	195	0.026		0.185	195	0.139	195

Class type	Father has no degree	Obs	Father had vocational training	Obs	Father has secondary certificate	Obs	Father has high school degree	Obs	Father has a college degree	Obs
Male math teacher	–	–	0.632	87	0.000	87	0.103	87	0.264	87
Female math teacher	–	–	0.654	107	0.009	107	0.112	107	0.224	107
Difference	–	–	−0.022		−0.009		−0.009		0.040	
(*t*-statistic)			(−0.317)		(−0.901)		(−0.193)		(0.645)	
Total	–	–	0.644	194	0.005	194	0.108	194	0.242	194

[1]–[4] See notes for Table 10A. [5] The first group are those students who had a male math teacher throughout their school career up to the point in time when the survey was conducted. The second group includes those students who had a female math teacher during the same period. We have dropped 17 students from the sample who were taught both by male and female math teachers. [6] No test results can be reported for the variable "Father has no degree" as this case does not apply to any of the students who had a male or female math teacher throughout their time at this school.

* Significant at the 10 percent level.

Table 11
Summary statistics.

Variable		Mean	St. dev.	Min	Max	Observations
Math grade	Overall	4.547	0.725	1.5	6	$N = 3942$
	Between		0.623	2.083	6	$n = 808$
	Within		0.379	2.922	6.381	$T = 4.879$
German grade	Overall	4.794	0.428	3	6	$N = 3942$
	Between		0.320	3.75	5.7	$n = 808$
	Within		0.293	3.694	6.128	$T = 4.879$
Female student	Overall	0.832	0.374	0	1	$N = 3942$
	Between		0.379	0	1	$n = 808$
	Within		0	0.832	0.832	$T = 4.879$

Table 11 (*Continued*)

Variable		Mean	St. dev.	Min	Max	Observations
Age	Overall	17.622	1.348	15	25	$N = 3942$
	Between		0.962	15.5	24.5	$n = 808$
	Within		1.069	15.622	20.288	$T = 4.879$
Single-sex class	Overall	0.228	0.420	0	1	$N = 3942$
	Between		0.416	0	1	$n = 808$
	Within		0.014	−0.022	0.978	$T = 4.879$
Number of male students	Overall	3.789	2.266	0	8	$N = 3942$
	Between		2.258	0	7.667	$n = 808$
	Within		0.419	−0.711	5.456	$T = 4.879$
Class size	Overall	21.102	2.391	12	25	$N = 3942$
	Between		2.088	14.667	25	$n = 808$
	Within		1.083	17.936	25.602	$T = 4.879$
Female German teacher	Overall	0.208	0.406	0	1	$N = 3942$
	Between		0.348	0	1	$n = 808$
	Within		0.174	−0.458	1.042	$T = 4.879$
Female math teacher	Overall	0.382	0.486	0	1	$N = 3942$
	Between		0.446	0	1	$n = 808$
	Within		0.206	−0.285	1.215	$T = 4.879$

Notes: Report cards are handed out twice a year in the first two school years and only once a year at the end of the third and fourth school year. Hence, there are at most six observations for each student.

Appendix B. Additional results

See Tables 12–16.

Table 12
Effect of single-sex schooling on female students (mathematics), subsamples, ex-ante ability.

Subsample criteria Dependent variable: math grade	Full sample Model 19	Full sample Model 20	Full sample Model 21	Full sample Model 22	Female math teacher Model 23	Male math teacher Model 24
Math grade in qualifying exam	0.345***	0.346***	0.350***	0.313***	0.263***	0.343***
	(9.571)	(9.414)	(9.381)	(8.747)	(4.269)	(8.588)
Math grade in qual. exam × Single-sex class				0.138*	0.141**	0.216
				(1.677)	(2.043)	(1.608)
Single-sex class	0.139***	0.157***	0.302***	−0.368	−0.441*	−0.487
	(3.193)	(2.805)	(3.976)	(−1.130)	(−1.649)	(−0.918)
Age	−0.093***	−0.090***	−0.089***	−0.090***	−0.162***	−0.050*
	(−3.742)	(−3.573)	(−3.545)	(−3.591)	(−3.667)	(−1.854)
Class size	−0.010	−0.002	−0.002	−0.001	0.007	0.005
	(−1.245)	(−0.161)	(−0.155)	(−0.074)	(0.376)	(0.473)
Female math teacher		−0.003	0.036			
		(−0.038)	(0.535)			
Single-sex class × Female math teacher			−0.262**			
			(−2.570)			
Cohort fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
School year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.183	0.194	0.194	0.200	0.177	0.233
Observations	2454	2454	2454	2454	987	1467
Number of female students	497	497	497	497	267	330

Notes: *t*-statistics are in parentheses. Standard errors are robust to heteroscedasticity and clustered at the class level. In models 23 and 24, the sum of female students is larger than 497 since some students were taught both by male and female math teachers.

* Significant at the 10 percent level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level.

Table 13
Effect of single-sex schooling on female students, inclusion of male students (grading-on-a-curve).

Dependent variable:	Math grade		German grade	
	Model 25	Model 26	Model 27	Model 28
Single-sex class	0.151***	0.166***	0.020	0.001
	(2.784)	(3.141)	(0.913)	(0.028)
Age	−0.072***	−0.073***	−0.028**	−0.028**
	(−3.555)	(−3.554)	(−2.246)	(−2.181)
Class size	0.010	0.010	−0.003	−0.003
	(1.090)	(1.081)	(−0.441)	(−0.413)
Female teacher	0.002	0.002	0.005	0.004
	(0.035)	(0.039)	(0.120)	(0.108)
Female student		−0.071		0.087***
		(−1.160)		(3.409)
Cohort fixed effects	Yes	Yes	Yes	Yes
School year fixed effects	Yes	Yes	Yes	Yes
R^2	0.033	0.034	0.034	0.039
Observations	3942	3942	3942	3942
Number of students	808	808	808	808

Notes: *t*-statistics are in parentheses. Standard errors are robust to heteroscedasticity and clustered at the class level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level.

Table 14
Effect of single-sex schooling on female students (mathematics), subsamples, only first-year data.

Subsample criteria	Full sample	Full sample	Female math teacher	Male math teacher	Full sample	Student in single-sex class	Student in coed class
Dependent variable: math grade	Model 29	Model 30	Model 31	Model 32	Model 33	Model 34	Model 35
Single-sex class	0.115**	0.154**	0.141**	0.183*	0.195**		
	(2.175)	(2.208)	(2.572)	(1.945)	(2.208)		
Age	−0.101***	−0.101***	−0.169***	−0.048	−0.100***	−0.154**	−0.082**
	(−3.079)	(−3.077)	(−3.354)	(−1.228)	(−3.063)	(−2.304)	(−2.243)
Female math teacher		−0.101			−0.082	−0.161***	−0.091
		(−1.144)			(−0.821)	(−9.030)	(−0.867)
Class size	0.010	0.009	0.055**	−0.022	0.010	0.019	0.002
	(0.543)	(0.502)	(1.991)	(−0.834)	(0.578)	(0.519)	(0.084)
Single-sex class × Female teacher					−0.078		
					(−0.708)		
Cohort fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1240	1240	433	807	1240	328	912
Number of female students	627	627	219	409	627	166	462

Notes: *t*-statistics are in parentheses. Standard errors are robust to heteroscedasticity and clustered at the class level. In models 29 and 30, the sum of female students is larger than 627 since some students were taught by both male and female teachers. In addition, with regard to models 32 and 33 there is one student who switched classes.
* Significant at the 10 percent level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level.

Table 15
Effect of single-sex schooling on female students (mathematics), excluding one math teacher at a time.

Dependent variable: math grade	Without math teacher #1 Model 4a	Without math teacher #2 Model 4b	Without math teacher #3 Model 4c	Without math teacher #4 Model 4d	Without math teacher #5 Model 4e	Without math teacher #6 Model 4f	Without math teacher #7 Model 4g	Without math teacher #8 Model 4h
Single-sex class	0.171***	0.218***	0.146***	0.165**	0.164***	0.193***	0.215***	0.135**
	(3.711)	(3.893)	(3.079)	(2.387)	(3.043)	(3.328)	(3.866)	(2.222)
Age	−0.088***	−0.075***	−0.060***	−0.060**	−0.070***	−0.067***	−0.071***	−0.086***
	(−3.190)	(−3.138)	(−2.617)	(−2.248)	(−3.153)	(−2.791)	(−3.277)	(−3.497)
Class size	0.007	0.009	0.009	0.006	0.008	0.004	0.014	0.007
	(0.693)	(0.933)	(0.881)	(0.676)	(0.818)	(0.435)	(1.274)	(0.660)
Female math teacher	0.030	−0.025	−0.068	0.037	−0.022	−0.032	−0.081	−0.021
	(0.414)	(−0.350)	(−1.522)	(0.475)	(−0.321)	(−0.438)	(−1.196)	(−0.297)
Cohort fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2566	2965	2897	2574	3139	3111	2842	2873
Number of classes	33	36	39	33	38	39	35	35

Notes: *t*-statistics are in parentheses. Standard errors are robust to heteroscedasticity and clustered at the teacher level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level.

Table 16
Effect of single-sex schooling on female students (mathematics), excluding one single-sex class at a time.

Dependent variable: math grade	Without single-sex class #1 Model 4i	Without single-sex class #2 Model 4j	Without single-sex class #3 Model 4k	Without single-sex class #4 Model 4l	Without single-sex class #5 Model 4m	Without single-sex class #6 Model 4n	Without single-sex class #7 Model 4o	Without single-sex class #8 Model 4p	Without single-sex class #9 Model 4q
Single-sex class	0.170***	0.135**	0.134**	0.204***	0.191***	0.177***	0.172***	0.166***	0.176***
	(2.863)	(2.475)	(2.249)	(3.633)	(3.502)	(2.850)	(3.134)	(3.043)	(3.224)
Age	−0.074***	−0.069***	−0.076***	−0.074***	−0.066***	−0.071***	−0.069***	−0.062***	−0.068***
	(−3.224)	(−2.908)	(−3.285)	(−3.183)	(−2.848)	(−3.011)	(−3.017)	(−2.769)	(−2.934)
Class size	0.006	0.007	0.006	0.007	0.007	0.007	0.009	0.004	0.009
	(0.635)	(0.729)	(0.562)	(0.763)	(0.790)	(0.743)	(0.956)	(0.394)	(0.870)
Female math teacher	−0.019	0.028	−0.011	−0.027	−0.015	−0.013	−0.018	−0.019	−0.016
	(−0.284)	(0.510)	(−0.156)	(−0.391)	(−0.229)	(−0.191)	(−0.267)	(−0.291)	(−0.241)
Cohort fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3171	3146	3157	3176	3171	3172	3211	3209	3219
Number of classes	38	38	38	38	38	38	38	38	38

Notes: *t*-statistics are in parentheses. Standard errors are robust to heteroscedasticity and clustered at the teacher level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level.

References

Ai, C., Norton, E.C., 2003. Interaction terms in logit and probit models. Econ. Lett. 80 (1), 123–129.
Arnott, R., Rowse, J., 1987. Peer group effects and educational attainment. J. Public Econ. 32, 287–305.
Beilock, S.L., Gunderson, E.A., Ramirez, G., Levine, S.C., 2010. Female teachers' math anxiety affects girls' math achievement. Proc. Natl. Acad. Sci. U.S.A. 107 (5), 1860–1863.
Black, S.E., Devereux, P.J., Salvanes, K.G., 2013. Under pressure? The effect of peers on outcomes of young adults. J. Labor Econ. 31 (1), 119–153.
Booth, A., Nolen, P., 2012a. Choosing to compete: how different are girls and boys? J. Econ. Behav. Organ. 81 (2), 542–555.
Booth, A.L., Nolen, P., 2012b. Gender differences in risk behaviour: does nurture matter? Econ. J. 122 (558), F56–F78.
Booth, A.L., Cardona-Sosa, L., Nolen, P., 2013. Do single-sex classes affect achievement? A study in a coeducational university. Banco de la Republica de Colombia.
Borghans, L., Meijers, H., Ter Weel, B., 2008. The role of noncognitive skills in explaining cognitive test scores. Econ. Inquiry 46 (1), 2–12.
Cameron, A.C., Gelbach, J.B., Miller, D.L., 2011. Robust inference with multiway clustering. J. Bus. Econ. Stat. 29 (2), 238–249.
Carrell, S.E., Page, M.E., West, J.E., 2010. Sex and science: how professor gender perpetuates the gender gap. Q. J. Econ. 125 (3), 1101–1144.
Compte, O., Postlewaite, A., 2004. Confidence-enhanced performance. Am. Econ. Rev. 94 (5), 1536–1557.
Cooray, A., Potrafke, N., 2011. Gender inequality in education: political institutions or culture and religion? Eur. J. Polit. Econ. 27 (2), 268–280.
Dee, T.S., 2005. A teacher like me: does race ethnicity, or gender matter? Am. Econ. Rev. 95 (2), 158–165.
Eisenkopf, G., 2010. Peer effects, motivation, and learning. Econ. Educ. Rev. 29 (3), 364–374.
Else-Quest, N.M., Hyde, J.S., Linn, M.C., 2010. Cross-national patterns of gender differences in mathematics: a meta-analysis. Psychol. Bull. 136 (1), 103–127.
Ferla, J., Valcke, M., Cai, Y., 2009. Academic self-efficacy and academic self-concept: reconsidering structural relationships. Learn. Individ. Differ. 19 (4), 499–505.
Foster, G., 2006. It's not your peers and it's not your friends. Some progress towards understanding the educational peer effect mechanism. J. Public Econ. 90, 1455–1475.
Fryer, R.G., Levitt, S.D., 2010. An empirical analysis of the gender gap in mathematics. Am. Econ. J. Appl. Econ. 2 (2), 210–240.

Gneezy, U., Rustichini, A., 2004. Gender and competition at a young age. Am. Econ. Rev. 94 (2), 377–381.
Gneezy, U., Niederle, M., Rustichini, A., 2003. Performance in competitive environments: gender differences. Q. J. Econ. 118 (3), 1049–1074.
Gneezy, U., Leonard, K.L., List, J.A., 2009. Gender differences in competition: evidence from a matrilineal and a patriarchal society. Econometrica 77 (5), 1637–1664.
Guiso, L., Monte, F., Sapienza, P., Zingales, L., 2008. Culture, gender, and math. Science 320 (5880), 1164–1165.
Gupta, N.D., Poulsen, A., Villeval, M.C., 2005. Male and female competitive behavior-experimental evidence. IZA Discussion Paper No. 1833.
Gupta, N.D., Poulsen, A., Villeval, M.C., 2013. Gender matching and competitiveness: Experimental evidence. Econ. Inquiry 51 (1), 816–835.
Halpern, D.F., Eliot, L., Bigler, R.S., Fabes, R.A., Hanish, L.D., Hyde, J., Liben, L.S., Martin, C.L., 2011. The pseudoscience of single-sex schooling. Science 333 (6050), 1706–1707.
Hoxby, C.M., 2000. The effects of class size on student achievement: new evidence from population variation. Q. J. Econ. 115 (4), 1239–1285.
Jackson, C.K., 2012. Single-sex schools, student achievement, and course selection: evidence from rule-based student assignments in Trinidad and Tobago. J. Public Econ. 96 (1–2), 173–187.
Kessels, U., Hannover, B., 2008. When being a girl matters less: accessibility of gender-related self-knowledge in single-sex and coeducational classes and its impact on students' physics-related self-concept of ability. Br. J. Educ. Psychol. 78 (2), 273–289.
Köller, O., Daniels, Z., Schnabel, K.U., Baumert, J., 2000. Kurswahlen von Mädchen und Jungen im Fach Mathematik: Zur Rolle von fachspezifischem Selbstkonzept und Interesse. Z. Padagog. Psychol. 14 (1), 26–37.
Köller, O., Baumert, J., Schnabel, K., 2001. Does interest matter? The relationship between academic interest and achievement in mathematics. J. Res. Math. Educ. 32 (5), 448–470.
Lavy, V., 2013. Gender differences in market competitiveness in a real workplace: evidence from performance-based pay tournaments among teachers. Econ. J. 123 (569), 540–573.
Lavy, V., Schlosser, A., 2011. Mechanisms and impacts of gender peer effects at school. Am. Econ. J. Appl. Econ. 3 (2), 1–33.
Lazear, E.P., 2001. Educational production. Q. J. Econ. 116 (3), 777–803.
Lee, V.E., Bryk, A.S., 1986. Effects of single-sex secondary schools on student achievement and attitudes. J. Educ. Psychol. 78 (5), 381–395.
Lee, V.E., Lockheed, M.E., 1990. The effects of single-sex schooling on achievement and attitudes in Nigeria. Comp. Educ. Rev. 34 (2), 209–231.
Lefgren, L., 2004. Educational peer effects and the Chicago public schools. J. Urban Econ. 56 (2), 169–191.
Link, C.R., Mulligan, J.G., 1991. Classmates' effects on black student achievement in public school classrooms. Econ. Educ. Rev. 10 (4), 297–310.
Manski, C.F., 1993. Identification of endogenous social effects: the reflection problem. Rev. Econ. Stud. 116 (2), 607–654.
Manski, C.F., 2000. Economic analysis of social interactions. J. Econ. Perspect. 13 (3), 115–136.
Marsh, H.W., Rowe, K.J., 1996. The negative effects of school-average ability on academic self-concept: an application of multilevel modelling. Aust. J. Educ. 40 (1), 65–87.
McEwan, P., 2003. Peer effects on student achievement: evidence from Chile. Econ. Educ. Rev. 22 (2), 131–141.
Neidell, M., Waldfogel, J., 2010. Cognitive and noncognitive peer effects in early education. Rev. Econ. Stat. 92 (3), 562–576.
Niederle, M., Vesterlund, L., 2007. Do women shy away from competition? Do men compete too much? Q. J. Econ. 122 (3), 1067–1101.
Niederle, M., Vesterlund, L., 2010. Explaining the gender gap in math test scores: the role of competition. J. Econ. Perspect. 24 (2), 129–144.
Oosterbeek, H., van Ewijk, R., 2014. Gender peer effects in university: evidence from a randomized experiment. Econ. Educ. Rev. 38, 51–63.
Ouazad, A., 2014. Assessed by a teacher like me: race, gender, and subjective evaluations. Educ. Finance Policy 9 (3), 334–372.
Park, H., Behrman, J.R., Choi, J., 2012. Causal effects of single-sex schools on college entrance exams and college attendance: random assignment in Seoul high schools. Demography 50 (2), 447–469.
Rockoff, J., 2009. Field experiments in class size from the early twentieth century. J. Econ. Perspect. 23 (4), 211–230.
Rotter, J., 1966. Generalized expectancies for internal versus external control of reinforcement. Psychol. Monogr. 80, 1–28.
Schneeweis, N., Zweimüller, M., 2012. Girls, girls, girls: gender composition and female school choice. Econ. Educ. Rev. 31 (4), 482–500.
Schütz, G., Ursprung, H., Wößmann, L., 2008. Education policy and equality of opportunity. Kyklos 61 (2), 279–308.
Shurchkov, O., 2012. Under pressure: gender differences in output quality and quantity under competition and time constraints. J. Eur. Econ. Assoc. 10 (5), 1189–1213.
Smith, P.B., Dugan, S., Trompenaars, F., 1997. Locus of control and affectivity by gender and occupational status: a 14-nation study? Sex Roles 36 (1–2), 51–77.
Smithers, A., Robinson, P., 2006. The paradox of single-sex and co-educational schooling. University of Buckingham, Buckingham.
Steele, C.M., 1997. A threat in the air. How stereotypes shape intellectual identity and performance. Am. Psychol. 52 (6), 613.
Sutter, M., Rützler, D., 2010. Gender differences in competition emerge early in life. University of Innsbruck. Working Papers in Economics and Statistics 14.
Thomson, T., Ungerleider, C., 2004. Single sex schooling: Final report – Canadian Centre for Knowledge Mobilisation. CCKM, Waterloo, Ontario.
Trautwein, U., Lüdtke, O., Köller, O., Baumert, J., 2006a. Self-esteem, academic self-concept, and achievement: how the learning environment moderates the dynamics of self-concept. J. Pers. Social Psychol. 90 (2), 334–349.
Trautwein, U., Ludtke, O., Marsh, H.W., Koller, O., Baumert, J., 2006b. Tracking, grading, and student motivation: using group composition and status to predict self-concept and interest in ninth-grade mathematics. J. Educ. Psychol. 98 (4), 788.
Trautwein, U., Lüdtke, O., Marsh, H.W., Nagy, G., 2009. Within-school social comparison: how students perceive the standing of their class predicts academic self-concept. J. Educ. Psychol. 101 (4), 853–866.
Whitmore, D., 2005. Resource and peer impacts on girls' academic achievement: evidence from a randomized experiment. Am. Econ. Rev. 95, 199–203.
Wigfield, A., Karpathian, M., 1991. Who am I and what can I do? Children's self-concepts and motivation in achievement situations. Educ. Psychol. 26 (3–4), 233–261.
Wößmann, L., West, M., 2006. Class-size effects in school systems around the world: evidence from between-grade variation in TIMSS. Eur. Econ. Rev. 50 (3), 695–736.