Academic Performance and Single-Sex Schooling: Evidence from a Natural Experiment in Switzerland

By GERALD EISENKOPF, ZOHAL HESSAMI, URS FISCHBACHER, AND HEINRICH W. URSPRUNG*

We study the effects of random assignment to coeducational and single-sex classes on the academic performance of female high school students. Our estimation results show that single-sex schooling improves the performance of female students in mathematics. This positive effect increases if the single-sex class is taught by a male teacher. 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. Single-sex schooling thus has profound implications for human capital formation and the mind-set of female students. (JEL I21, J16*)*

Gender gaps in academic performance, especially in mathematics, continue to be observed worldwide (Guiso, Monte, Sapienza, and Zingales, 2008; Else-Quest, Hyde, and Linn, 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

*Eisenkopf: Department of Economics, University of Konstanz (and Thurgau Institute of Economics, Kreuzlingen, Switzerland), Box 131, 78457 Konstanz, Germany (e-mail: Gerald.Eisenkopf@uni-konstanz.de); Hessami: Department of Economics, University of Konstanz, Box 138, 78457 Konstanz, Germany (e-mail: Zohal.Hessami@uni-konstanz.de); Fischbacher: Department of Economics, University of Konstanz (and Thurgau Institute of Economics, Kreuzlingen, Switzerland), Box 131, 78457 Konstanz, Germany (e-mail: Urs.Fischbacher@uni-konstanz.de); Ursprung: Department of Economics, University of Konstanz, Box 138, 78457 Konstanz, Germany (e-mail: Heinrich.Ursprung@uni-konstanz.de). The authors thank Lorenz Zubler, Georg Leumann, Claudia Peyer and Constanze Bürki from the Pädagogische Maturitätsschule Kreuzlingen for providing us with anonymous data from 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 and the ZEW Mannheim as well as conferences in Marbach (Swiss Leading House Economics of Education) and Tübingen (Economics of Education Commitee of the German Economic Association). Arye Hillman provided helpful comments. Christoph Kappeler and Anne Wunderlich provided excellent research assistance. Finally, we gratefully acknowledge financial support from the Young Scholar Fund of the University of Konstanz.

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.[1] So far, the gender composition of peer groups has received little attention in empirical education economics which can be attributed 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.

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.[2] 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. Moreover, in the Nigerian all-girls schools, which represent a subset of the overall sample, mathematics teachers happen to be exclusively female, implying that in this particular study gender-specific peer effects cannot be isolated from a potential indirect peer effect working through the teacher's gender.

In an influential recent contribution, Carell, Page, and West (2010) circumvent this problem by using a sample that includes both male and female instructors. Their estimation results suggest powerful professor gender effects on female students' performance in mathematics and science. Interestingly, these effects were identified in a coeducational environment at the college level (United States Air Force Academy). Based on these findings one could

[1] Goldin and Katz (2010) analyze the timing of coeducation in U.S. higher education and its consequences for women's general educational attainment.

[2] 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).

argue that similar effects may be at work in a high school environment with younger and more impressionable students.

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 on the academic performance of female students. Our identification strategy exploits a natural experiment at an upper-secondary high school in Switzerland, where the school board randomly assigns incoming female students to coeducational and single-sex classes. 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 self-selection problems at the class-level can be ruled out.

In addition to the gender of classmates, we are also interested in the impact of teacher gender. The reason is that peer effects depend on how the teacher relates to his or her class, and this relationship is in turn likely to be influenced by the gender composition of the class. In line with the existing literature on gender-specific peer effects on academic performance our main focus is on mathematics achievement. However, in order to allow for a comparison, we also use data on German language skills.

The estimation results indicate that gender-specific peer group effects are indeed at work: we find a positive effect of single-sex education on female students' proficiency in mathematics but not in German. In addition, the effect in mathematics tends to be stronger if female students in a single-sex class are taught by a male teacher.

We argue that the identified influence of single-sex education on mathematics achievement relies, at least to some extent, on a channel of influence running from the single-sex environment to the female student's mindset which, in turn, facilitates higher academic achievements. To support our conjecture, we conducted a questionnaire survey. The responses indicate that 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 German we do not observe these differences. This is a clear indication for a specific kind of social learning in a single-sex environment.

The remainder of the study is structured as follows: Section I provides a brief survey of related strands of the literature and advances three hypotheses. Section II describes the design of the study and the collected data. Section III elaborates on the empirical strategy, presents descriptive statistics, and reports the regression results. Section IV presents the survey-based evidence, and Section V offers some conclusions.

I. 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 five 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 three 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 first hypothesis on gender-specific (direct) peer effects. We then turn to the literature on teacher characteristics as a fourth potential source of the gender gap in mathematics and advance two additional hypotheses on the role of indirect peer effects and their interaction with direct peer effects.

A. The Math Gender Gap

The gender gap in mathematics has recently attracted a great deal of attention in education economics. Guiso, Monte, Sapienza, and Zingales (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.[3] In more gender-equal societies the math gender gap turns out to be smaller; the gap even disappears in countries enjoying very high gender equality, such as Norway and Sweden.

In a large panel data set which is 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 strata 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 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.

B. Explaining the Math Gender Gap

Gender Differences in Competitiveness.—A large body of literature establishes that men are in general more willing to compete than women (Gupta, Poulsen, and Villeval, 2005; Niederle and Vesterlund, 2007; Gneezy, Leonard, and List, 2009). Sutter and Rützler (2010) even find a gender gap in

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

competitiveness among three-year olds. Since it is conceivable that standard math tests take place in a more competitive environment than, for example, writing essays, more competitive students have an advantage in math tests. Thus, 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. While Gupta et al. (2005) indeed provide evidence for a significant effect of the competitor's gender in a series of experiments, Gneezy, Niederle and Rustichini (2003) and Gneezy and Rustichini (2004) find no effect. 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 competitors' gender is that such effects may relate to the gender composition of the environment in which the female students are educated. In a field experiment, Booth and Nolen (2009a; 2009b) examine this question with regard to students just under 15 years of age attending publicly-funded single-sex and coeducational schools. 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, which implies that the gender composition in the classroom has a persistent effect on female students' competitive behavior.

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.

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 way students perceive themselves and their achievements denoted in the psychological literature as the *locus of control* which can be either internal or external. 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, Dugan, and Trompenaars, 1997), female students may be more easily distracted 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, maybe because failing in mathematics or physics can no longer be explained away by claiming that male classmates have an innate advantage.

A second important dimension of students' self-perception is the so-called *academic self-concept* which refers to students' self-perceptions regarding their academic achievements (Wigfield and Karpathian, 1991; Ferla, Valcke, and Cai, 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, Baumert and Schnabel (2001), for example, 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, Lüdtke, Köller, and Baumert, 2006a; Trautwein et al., 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, Lüdtke, Marsh, and Nagy, 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.

Peer-Group Effects.—Analyzing peer-group effects has a long tradition in education economics. Neidell and Waldfogel (2010) find that the unruly behavior of children with limited self-control or discipline has a bad influence on their peers. Other studies investigate how educational outcomes are affected by 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).

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 (2009) succeed in identifying a causal impact of the gender composition in coeducational classes on female students' choice of secondary school type. By exploiting the natural variation in the gender composition of adjacent cohorts within schools, they show that girls are less likely to choose female-dominated types of secondary schools and more likely to choose a technical school type if they were exposed to a larger share of female peers in previous grades.

In combination, these studies lead us to conjecture that the gender composition in the learning environment will have immediate effects on the academic performance of female students and circuitous effects working through the differential acquisition of non-cognitive skills. We thus propose:

Hypothesis 1: Classmate Gender Effects
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 literature has hitherto focused on the math gender gap. The three potential explanations for the math gender gap may, of course, also apply to other academic subjects. We conjecture, however, that the respective channels of influence are not at work in subjects in which female students on average do not perform worse than male students. To examine this conjecture we also investigate in Sections III and IV the influence of single-sex schooling on German language skills.

Teacher Characteristics.—A large body of literature analyzes the impact of various teacher characteristics on student achievement. The best researched characteristic is teacher quality even though teacher quality is extremely hard to measure. Usually, it is either proxied by experience and training, or it is determined as a residual of regression estimates. It transpires that the influence of teacher quality is not as clear-cut as one may expect. To be sure, some studies do find a positive impact of teacher quality on student achievement (Rivkin, Hanushek and Kain, 2005; Rockoff, 2004). On the other hand, Jepsen (2005) uses survey data and finds that a number of non-standard teacher characteristics are insignificant predictors of student achievement, especially so in lower grades.

In addition, a number of studies investigate whether racial, ethnic, and gender matching of students and teachers influence the students' academic achievements. While the studies by Dee (2004; 2005) find significant and large effects for all three dimensions of student-teacher matching, Howsen and Trawick (2007) provide evidence that race matching of teachers and students has no statistically significant effect if one controls for student innate ability and teacher gender.

Lavy (2008) focuses on the interaction of student and teacher gender in teacher's evaluations of students' performance. He uses a natural experiment performed at the academic track of Israeli high schools where the matriculation examination comes in two parts, both having the same structure: an anonymous all-state test and a school-level test graded by the student's teacher who, of course, is aware of the student's gender. Comparing the two test components, Lavy finds that male students face discrimination in all segments of the ability and performance distribution. Since the size of the difference between the two test results is very sensitive to the teachers' characteristics, the bias against male students appears to be the result of teachers', and not students', behavior.

In a similar attempt, Holmlund and Sund (2008) use data from upper-secondary schools in Stockholm to investigate whether the observed superior academic performance of girls can be attributed to the female domination in the teacher profession. They find that gender-specific performance differences indeed increase with the share of female teachers. This effect can, however, not be interpreted as being causal because of nonrandom assignment of teachers to classrooms. Moreover, the authors do not find strong support for their initial hypothesis when they relate changes in student performance to reassignments of students to teachers of the same sex.

The fact that teacher gender may have a very direct effect on student performance has already been pointed out in the introduction. The study by Carrell et al. (2010) makes use of a random assignment of students to classes at the US Air Force Academy where all students take the same exam. Controlling for student ability as measured by SAT scores, two thirds of the gender gap in grade points disappears when a female professor teaches a mathematics or science class. On the other hand, professor gender has little impact on male students, while top-performing female students benefit the most from female professors.

Based on this extensive literature we put forward the following hypothesis:

Hypothesis 2: Teacher Gender Effect

The academic performance of female students in mathematics is influenced by teacher gender.

The effect described in hypothesis 2 may be due to teachers (subconsciously) discriminating either against boys or girls, or it may arise because the teacher develops a symbiotic relationship with the class. The latter effect is especially plausible if all students are female. Therefore, our third hypothesis captures the fact that teacher-student interactions may depend on the gender composition of a class:

Hypothesis 3: Teacher Gender – Peer Gender Interaction

The influence of teacher gender on the academic performance of female students in mathematics differs between single-sex and coeducational classes.

II. Data

Our study was conducted at a Swiss high school.[4] The four-year curriculum of this school prepares the students to obtaining the matriculation certificate. Catering especially to students who, upon graduation, aspire to attend a college of education, the school's curriculum places emphasis on pedagogical subjects, but covers all the basic subjects required at upper-secondary schools in Switzerland. The focus on teacher training allows the school's graduates to skip the basic first-year courses at the University of Teacher Education located in the same town. This special 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. The school, in particular, does not allow for self-selection. The assignment is thus based on a real random process.

We have culled our key data from the school's administrative records. These records contain information on all students who have attended the school from the school years 2001/02 to 2008/09.[5] Our dataset comprises 808 students for whom we have information on characteristics such as gender, date of birth, classmates, and report card grades. In each school year, there are four to five classes with about 18 to 25 students per class. Each student takes some 12 to 13 courses. Both German and math are compulsory. Table 1 describes the composition of the sample.

---

[4] Pädagogische Maturitätsschule (PMS) Kreuzlingen

[5] 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

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.

Our data also 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.

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	19.4

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

We also obtained data detailing the incoming students' performance in the entry exams. Students typically have to pass a written exam before enrolling at an upper-secondary school in their respective home canton. Students can take the exam at different locations and an overall passing grade allows them to attend any upper-secondary school in the canton. We obtained the entry exam grades for most students (599 out of 808), where earlier cohorts are excluded because of changes in the admission and examination policies. These entry exam grades provide a standardized measure of 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.

Table 3 indicates a rather small grade difference in 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.

TABLE 3—SUMMARY OF QUALIFYING EXAM GRADES

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

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 10 to 12 in the appendix describe all of the variables that are included in the empirical analysis.

III. Empirical Results

A. Descriptive Analysis

We measure academic performance with ordinary report card grades because public schools in the canton of Thurgau do not run standardized end-of-school-year or exit exams. Grades are, however, a highly incentivized measure as they determine at the end of each school year 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 these criteria very conscientiously since they teach both types of classes.[6] 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.

Figure 1 shows the distribution of math and German grades of female students. The best grade that can be achieved is 6. Grades decline in steps of 0.5 and 1 is the worst possible grade. A grade of 4 is the minimum grade required to pass an exam and to avoid retention. 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.

[6] 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.

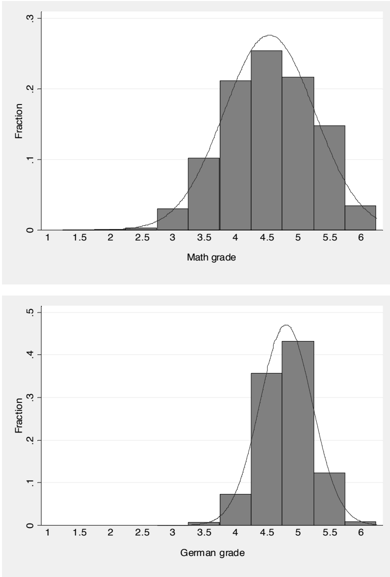

FIGURE 1. DISTRIBUTION OF MATHEMATICS AND GERMAN GRADES ACROSS FEMALE STUDENTS

B. Identification strategy

Our empirical model has the following form:

$$\begin{aligned} Grade_{itc} = \alpha_i + \beta\mathbf{X}_{itc} + \delta Singlesexclass_{itc} + \mu Femaleteacher_{tc} \\ + \varphi Schoolyear_{tc} + \zeta Singlesexclass_{itc}{*}Schoolyear_{tc} \\ + \chi Singlesexclass_{itc}{*}Femaleteacher_{tc} + \omega_c + \varepsilon_{itc}, \end{aligned}$$

where the dependent variable measures the report card grade in either mathematics or German. Subscript i denotes the respective female student (i = 1, 2, …, 668), c the respective class (c = 1, 2, …, 39), and t the respective report card (t = 1,2,…,6).

The vector $\mathbf{X}_{itc}$ includes the age of the student and the number of students in his or her class, while the cohort dummies ω_c capture the fact that some cohorts may be more proficient in mathematics or German than others. With regard to the hypotheses outlined in section I, δ is relevant for hypothesis 1, μ for hypothesis 2, and χ for hypothesis 3. The estimate of ζ provides information on whether single-sex schooling, if it has an effect at all, immediately impacts on the female students' performance or whether this effect emerges only after some extended time of single-sex schooling.

To estimate equation (1), we rely on OLS estimations with random effects at the student-level, and robust standard errors clustered at the class-level. 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 run 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 the next section with an ordered probit estimator and briefly discuss the marginal effects of all relevant coefficients except those of the interaction terms.[7]

C. Estimation Results

Our 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. Since virtually all grades range between 3.5 and 6, a coefficient of 0.25 implies a performance

[7] 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 OLS. Figure 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).

increase of about 10% of the relevant range.[8] The math teacher dummies in model 1 turn out to be jointly significant (the p-value corresponds to significance at the 5% level), indicating that grading comprises an idiosyncratic element. Model 2 includes a female teacher dummy instead of dummies for each teacher as in model 1. The estimation results reject hypothesis 2. Hence, it is not the teacher's gender that is driving the teacher-specific grading differences.

TABLE 4— RANDOM EFFECTS ESTIMATIONS: FEMALE STUDENTS (MATHEMATICS)

Dependent variable: math grade	Model 1	Model 2	Model 3
Single-sex class	0.232***	0.172***	0.211***
	(3.822)	(3.137)	(2.634)
School year	0.105***	0.109***	0.102***
	(4.361)	(4.584)	(4.008)
Age	-0.069***	-0.070***	-0.068***
	(-3.034)	(-3.115)	(-3.030)
Female math teacher		-0.016	
		(-0.250)	
Class size	0.009	0.008	0.009
	(0.941)	(0.861)	(0.963)
Single-sex class * School year			0.010
			(0.489)
Math teacher dummies	YES	NO	YES
p-value for joint significance of teacher dummies	0.043	-------	0.049
R^2	0.039	0.040	0.040
Observations	3,281	3,281	3,281
Number of female students	668	668	668

Notes: All estimations include dummies for the eight different cohorts. t-statistics are in parentheses. Standard errors are robust to heteroscedasticity and clustered at the class-level.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

In model 3, we include an interaction term between class type and school year, testing whether the class type effect increases over time. This is not the case. We find no evidence that the positive single-sex class effect on female students' performance significantly increases as the students advance to higher grades.

[8] In the corresponding ordered probit estimations of models 1 and 2, we find that female students in single-sex math classes have a 2.32-2.61% higher probability of obtaining the highest grade of 6 than female students in coed classes. This effect is significant at the 10 percent level for both models. In comparison, each year of age reduces the probability of obtaining the highest grade in math by 1.6%, while advancing to a higher school year increases this probability by 2.25%. These two marginal effects are also significant at the 10 percent level.

The coefficient estimates of the remaining control variables included in the regressions illustrate that students perform better as they advance from 9th to 12th grade, whereas older students of a given cohort perform worse. Both effects are driven by a retention policy that forces poorly performing students to repeat a school year. Class size does not appear to have an influence on academic achievements.[9]

As already pointed out, model 2 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 4 and 5 reveal that single-sex schooling benefits female students regardless of teacher gender. However, the effect is smaller for female teachers. In model 6, we report results for a regression containing 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 gender-specific peer effect. In models 7 and 8, 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. The high t-stastistic of -12.375 is particularly noteworthy and shows the robustness of the effect. 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. This result supports hypothesis 3.

We also tested whether the academic performance of boys or girls increases if the number of male students in a coeducational class gradually increases. Yet, we do not observe such an effect in our sample. We are therefore led to conclude that it is the very absence of male students that drives our results. At

---

[9] 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.

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 high school environment (in Switzerland only about 20% of the 15- to 18-year-olds attend high school) which, moreover, is dominated by students who aspire to become teachers themselves.

TABLE 5— RANDOM EFFECTS ESTIMATIONS: FEMALE STUDENTS (MATHEMATICS), SUBSAMPLES

Subsample criteria	Female math teacher	Male math teacher	Full sample	Student in single-sex class	Student in coed class
Dependent variable: math grade	Model 4	Model 5	Model 6	Model 7	Model 8
Single-sex class	0.138***	0.303***	0.319***		
	(3.544)	(4.023)	(4.030)		
School year	0.152***	0.084***	0.110***	0.152***	0.102***
	(3.576)	(2.947)	(4.638)	(3.356)	(3.552)
Age	-0.120***	-0.047*	-0.070***	-0.088**	-0.065**
	(-3.431)	(-1.682)	(-3.129)	(-2.263)	(-2.391)
Female math teacher			0.030	-0.389***	0.037
			(0.527)	(-12.375)	(0.652)
Class size	0.011	0.013	0.007	0.019	-0.000
	(0.683)	(0.947)	(0.787)	(1.420)	(-0.030)
Single-sex class * Female teacher			-0.269***		
			(-2.950)		
R^2	0.057	0.040	0.038	0.042	0.033
Observations	1,316	1,965	3,281	898	2,383
Number of female students	366	444	668	180	489

Notes: All estimations include dummies for the eight different cohorts. t-statistics are in parentheses. Standard errors are robust to heteroscedasticity and clustered at the class-level. In models 4 and 5, 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 7 and 8 there is one student who switched classes.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

We now turn to the analysis of the academic performance in German. Table 6 reports the results. The estimates indicate that in language arts (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.[10]

TABLE 6— RANDOM EFFECTS ESTIMATIONS: FEMALE STUDENTS (GERMAN)

Dependent variable: German grade	Model 9	Model 10	Model 11	Model 12
Single-sex class	-0.029	-0.012	-0.034	-0.032
	(-0.986)	(-0.552)	(-0.741)	(-1.163)
School year	0.094***	0.091***	0.093***	0.093***
	(4.532)	(4.385)	(4.286)	(4.404)
Age	-0.052***	-0.054***	-0.052***	-0.053***
	(-4.023)	(-4.085)	(-4.024)	(-4.006)
Female German teacher		0.024		-0.007
		(0.526)		(-0.114)
Class size	-0.012	-0.012	-0.012	-0.013
	(-1.489)	(-1.411)	(-1.486)	(-1.561)
Single-sex class * School year			0.002	
			(0.128)	
Single-sex class * Female teacher				0.078
				(1.105)
German teacher dummies	YES	NO	YES	NO
p-value for joint significance of German teacher dummies	0.000	-------	0.000	-------
R^2	0.019	0.014	0.019	0.014
Observations	3,281	3,281	3,281	3,281
Number of female students	668	668	668	668

Notes: All estimations include dummies for the eight different cohorts. t-statistics are in parentheses. Standard errors are robust to heteroscedasticity and clustered at the class-level.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

D. Additional Specifications

Innate Ability.—The results summarized in Table 7 illustrate that the estimates presented in the previous tables are independent of students' innate abilities as measured by the grades received in the qualifying exam. The four models in Table 7A correspond to models 1 to 3 in Table 4 and to model 6 in Table 5. The second set of four models (Table 7B) corresponds to models 9 to 12 in Table 6. The grades received in the qualifying exam have, not surprisingly, a strong explanatory power for the students' subsequent

[10] In the corresponding ordered probit estimations of models 9 and 10, we also find that the coefficient and the marginal effect of the single-sex class dummy is insignificant with t-statistics of -0.36 and-0.06, respectively.

academic performance and capture a substantial share of the ex-ante heterogeneity among the student body.

TABLE 7A— ROBUSTNESS CHECK I: CONTROLLING FOR QUALIFYING EXAM GRADES IN MATHEMATICS

Dependent variable:	Math grade			
	Model 13	Model 14	Model 15	Model 16
Math grade in qualifying exam	0.342***	0.346***	0.342***	0.350***
	(9.161)	(9.431)	(9.169)	(9.382)
Single-sex class	0.213***	0.156***	0.172**	0.296***
	(3.813)	(2.829)	(2.261)	(4.234)
School year	0.130***	0.140***	0.125***	0.140***
	(4.731)	(4.885)	(4.356)	(4.981)
Age	-0.088***	-0.091***	-0.088***	-0.090***
	(-3.521)	(-3.576)	(-3.506)	(-3.568)
Female Math teacher		-0.004		0.031
		(-0.062)		(0.460)
Class size	-0.001	-0.002	-0.001	-0.003
	(-0.098)	(-0.238)	(-0.106)	(-0.325)
Single-sex class * School year			0.020	
			(0.743)	
Single-sex class * Female math				-0.253***
				(-2.684)
Math teacher dummies	YES	NO	YES	NO
R^2	0.194	0.194	0.194	0.194
Observations	2,454	2,454	2,454	2,454
Number of female students	497	497	497	497

TABLE 7B— ROBUSTNESS CHECK I: CONTROLLING FOR QUALIFYING EXAM GRADES IN GERMAN

Dependent variable:	German grade			
	Model 17	Model 18	Model 19	Model 20
German grade in qualifying exam	0.153***	0.154***	0.153***	0.153***
	(6.454)	(6.529)	(6.447)	(6.368)
Single-sex class	-0.041	-0.013	-0.049	-0.023
	(-1.018)	(-0.473)	(-0.852)	(-0.713)
School year	0.099***	0.096***	0.098***	0.097***
	(4.493)	(4.391)	(4.315)	(4.434)
Age	-0.055***	-0.057***	-0.055***	-0.057***
	(-4.047)	(-4.107)	(-4.037)	(-4.068)
Female German teacher		-0.006		-0.019
		(-0.125)		(-0.308)
Class size	-0.013	-0.012	-0.013	-0.013
	(-1.375)	(-1.281)	(-1.379)	(-1.327)
Single-sex class * School year			0.004	
			(0.184)	
Single-sex class * Female German				0.042
				(0.608)
German teacher dummies	YES	NO	YES	NO
R^2	0.085	0.085	0.085	0.085
Observations	2,454	2,454	2,454	2,454
Number of female students	497	497	497	497

Notes: All estimations include dummies for the eight different cohorts. t-statistics are in parentheses. Standard errors are robust to heteroscedasticity and clustered at the class-level.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Grading on a Curve.—We also investigated whether grading-on-a-curve might be responsible for the significant single-sex class coefficient. Even though it would counteract the explicit school policy, it is not entirely inconceivable that some teachers may grade according to a predefined grade distribution that is imposed on each class. If boys performed better than girls, girls 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 boys present to capture the highest grades.

In Table 8, we report the results of four regressions that are based on the grades of male and female students. Models 21 and 22 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. With regard to the academic performance in German, the estimation results for models 23 and 24 do not indicate any differences between single-sex and coeducational classes.

TABLE 8— ROBUSTNESS CHECK II: INCLUSION OF MALE STUDENTS TO INVESTIGATE GRADING-ON-A-CURVE

Dependent variable:	Math grade		German grade	
	Model 21	Model 22	Model 23	Model 24
Single-sex class	0.210***	0.225***	-0.012	-0.030
	(3.28)	(3.61)	(-0.40)	(-0.98)
School year	0.097***	0.097***	0.082***	0.082***
	(4.12)	(4.11)	(4.60)	(4.54)
Age	-0.072***	-0.072***	-0.040***	-0.040***
	(-3.57)	(-3.56)	(-3.21)	(-3.13)
Class size	0.010	0.010	-0.008	-0.008
	(0.99)	(0.98)	(-1.14)	(-1.12)
Female student		-0.068		0.082***
		(-1.12)		(3.29)
Teacher dummies	YES	YES	YES	YES
R^2	0.030	0.032	0.020	0.024
Observations	3,942	3,942	3,942	3,942
Number of students	808	808	808	808

Notes: All estimations include dummies for the eight different cohorts. t-statistics are in parentheses. Standard errors are robust to heteroscedasticity and clustered at the class-level.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Overall, our results support our first hypothesis claiming that the academic performance of female students varies with the gender composition of their classmates. Two 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 in the sense that the mere presence of male students compromises the educational environment that is especially conducive to the female students' academic development.[11] Second, single-sex education is not advantageous to female students across the board: in some subjects (mathematics) the advantages are sizable; in other subjects (German) no significant effects can be identified.

Our hypothesis concerning the impact of teacher gender on gender-specific peer effects are also confirmed for math performance. We show that class type is crucial for the influence of teacher gender on female students' performance. Teacher gender influences the high school students' academic achievements only in single-sex classes and in specific subjects: In all-girls classes, male teachers are able to elicit better accomplishments in mathematics, but not in German.[12]

IV. 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, Baumert, and Schnabel, 2001; Trautwein et al., 2006a; Trautwein et al., 2006b; Trautwein et al., 2009). Because of this established link between self-perceived competence

[11] In our sample we have coeducational classes with 2 up to 8 male students.

[12] 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 Carell 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.

to 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.

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.

A by-product of our survey consists in our not finding any 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 with the girls' assignment to the two types of classes: the families of girls assigned to single-sex classes own fewer musical instruments and their mothers or fathers 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. If this were indeed the case, our results would actually be even stronger since it is well known that the socio-economic status has a detrimental effect on academic achievement (Schütz, Ursprung, and Wößmann, 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.[13] Not surprisingly, both measures turn out to be highly correlated with performance.

Tables 9A and 9B 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 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.[14]

TABLE 9A— 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
Class type						
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 9B— SURVEY RESPONSES BY FEMALE STUDENTS (9TH GRADE, I.E. COHORT 9)[15] ATTENDING SINGLE-SEX AND COEDUCATIONAL CLASSES IN SPRING 2010

	Math self-concept		Math self-assessment		German self-assessment	
	Observations	Response	Observations	Response	Observations	Response
Class type						
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

13 We use the same statements that have been used in relevant psychological studies in German speaking countries (e.g. Köller, Daniels, Schnabel, and Baumert, 2000; Köller et al., 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.

14 There is no single-sex class in the second year.

15 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.).

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 driver 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 2009) and lower levels of risk aversion (Booth and Nolen 2010). On a more methodological level, it is worthwhile to point out that our survey-based observations nicely back up our claim that the observed effect of single-sex education is not likely to be attributable to capricious grading. 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, 2009).

V. 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 girls. 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 problems relating to self-selection: 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 almost inexistent.

We provide the first evaluation of female single-sex education with a randomized assignment of girls into different learning environments and find strong empirical support for the benefits of single-sex education. Analyzing a natural experiment performed at an upper-secondary 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 German. Moreover, the effect in mathematics tends to be stronger if girls in a single-sex class are taught by a male teacher.

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 girls 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 result rules out that our findings are an artefact of an implicit grading-on-a-curve policy.

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 data from the survey study are in line with this interpretation. Whether the stereotype threat paradigm can explain the identified teacher-gender 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, Gunderson, Ramirez, and Levine, 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.

Our second hypothesis derives from the notion that peer-competition is a major driving force of 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 Figure 1 clearly 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. The less competitive girls, 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 will therefore spend more effort in single-sex classes and accordingly perform better. This is the direct peer effect of single-sex schooling.

Whether male teachers boost the competitive spirit in all-girls classes via a "groupie effect", as it were, and thereby provide an additional, albeit indirect, peer effect is pure speculation. In any event, we know from the research by Booth and Nolen (2009) that girls educated in a single-sex environment behave more like boys in competitive situations. Together with the direct peer effect which renders competition more rewarding per se, this socially acquired competitive spirit provides single-sex schooling with an additional advantage that makes itself be felt especially in 'male' subjects such as mathematics.

Even though the identified positive effect of single-sex schooling 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. Our second interpretation is admittedly purely ad hoc and might even appear rather frivolous to some readers. One advantage of such speculations is, however, to spur disagreement and future research.

APPENDIX

TABLE 10—DEFINITIONS OF VARIABLES

Variable	Description
Dependent variables	
Math grade	Respective grade in each report card measured on a scale from 1
German grade	(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)

TABLE 11—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 rightaway: "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 12—SUMMARY STATISTICS

Variable		Mean	Std.Dev.	Min	Max	Observations
Math grade	Overall	4.547	0.725	1.5	6	N = 3,942
	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 = 3,942
	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 = 3,942
	Between		0.379	0	1	n = 808
	Within		0	0.832	0.832	T = 4.879
Age	Overall	17.622	1.348	15	25	N = 3,942
	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 = 3,942
	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 = 3,942
	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 = 3,942
	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 = 3,942
	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 = 3,942
	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.

REFERENCES

Ai, C. and E. C. Norton (2003). "Interaction terms in logit and probit models." Economics Letters, 80(1): 123-29.

Arnott, R. and J. Rowse (1987). "Peer group effects and educational attainment." Journal of Public Economics, 32: 287-305.

Beilock, S. L., E. A. Gunderson, G. Ramirez and S. C. Levine (2010). "Female teachers' math anxiety affects girls' math achievement." Proceedings of the National Academy of Sciences, 107(5): 1860-63.

Booth, A. L. and P. J. Nolen (2009a). "Choosing to compete: How different are girls and boys?" CEPR Discussion Paper No. 7214.

Booth, A. L. and P. J. Nolen (2009b). "Gender differences in risk behaviour: Does nurture matter?" IZA Discussion Paper No. 4026.

Borghans, L., H. Meijers and B. ter Weel (2008). "The role of noncognitive skills in explaining cognitive test scores." Economic Inquiry, 46(1): 2-12.

Carrell, S. E., M. E. Page and J. E. West (2010). "Sex and science: How professor gender perpetuates the gender gap." Quarterly Journal of Economics, 125(3): 1101-44.

Compte, O. and A. Postlewaite (2004). "Confidence-enhanced performance." American Economic Review, 94(5): 1536-57.

Cooray, A. and N. Potrafke (2011). "Gender inequality in education: Political institutions or culture and religion?" European Journal of Political Economy, 27(2): 268-80.

Dee, T. S. (2004). "Are there civic returns to education?" Journal of Public Economics, 88(9-10): 1697-720.

Dee, T. S. (2005). "A teacher like me: Does race, ethnicity, or gender matter?" American Economic Review, 95(2): 158-65.

Eisenkopf, G. (2010). "Peer effects, motivation, and learning." Economics of Education Review, 29(3): 364-74.

Else-Quest, N. M., J. S. Hyde and M. C. Linn (2010). "Cross-national patterns of gender differences in mathematics: A meta-analysis." Psychological Bulletin, 136(1): 103-27.

Ferla, J., M. Valcke and Y. Cai (2009). "Academic self-efficacy and academic self-concept: Reconsidering structural relationships." Learning and Individual Differences, 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." Journal of Public Economics, 90: 1455-75.

Fryer Jr., R. G. and S. D. Levitt (2010). "An empirical analysis of the gender gap in mathematics." American Economic Journal: Applied Economics, 2(2): 210-40.

Gneezy, U., K. L. Leonard and J. A. List (2009). "Gender differences in competition: Evidence from a matrilineal and a patriarchal society." Econometrica, 77(5): 1637-64.

Gneezy, U., M. Niederle and A. Rustichini (2003). "Performance in competitive environments: Gender differences." Quarterly Journal of Economics, 118(3): 1049-74.

Gneezy, U. and A. Rustichini (2004). "Gender and competition at a young age." American Economic Review Papers and Proceedings, 94(2): 377-81.

Goldin, C. and L. F. Katz (2010). "Putting the co in education: Timing, reasons, and consequences of college coeducation from 1835 to the present" NBER Working Paper No. 16281.

Guiso, L., F. Monte, P. Sapienza and L. Zingales (2008). "Culture, gender, and math." Science, 320(5880): 1164-65.

Gupta, N. D., A. Poulsen and M. C. Villeval (2005). "Male and female competitive behavior-experimental evidence." IZA Discussion Paper No. 1833.

Holmlund, H. and K. Sund (2008). "Is the gender gap in school performance affected by the sex of the teacher?" Labour Economics, 15(1): 37-53.

Howsen, R. M. and M. W. Trawick (2007). "Teachers, race and student achievement revisited." Applied Economics Letters, 14(14): 1023-27.

Jepsen, C. (2005). "Teacher characteristics and student achievement: Evidence from teacher surveys." Journal of Urban Economics, 57(2): 302-19.

Kessels, U. and B. Hannover (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." British Journal of Educational Psychology, 78(2): 273-89.

Köller, O., J. Baumert and K. Schnabel (2001). "Does interest matter? The relationship between academic interest and achievement in mathematics." Journal for Research in Mathematics Education, 32(5): 448-70.

Köller, O., Z. Daniels, K. U. Schnabel and J. Baumert (2000). "Kurswahlen von Mädchen und Jungen im Fach Mathematik: Zur Rolle von fachspezifischem Selbstkonzept und Interesse." Zeitschrift für Pädagogische Psychologie, 14(1): 26-37.

Lavy, V. (2008). "Do gender stereotypes reduce girls' or boys' human capital outcomes? Evidence from a natural experiment." Journal of Public Economics, 92(10-11): 2083-105.

Lee, V. E. and A. S. Bryk (1986). "Effects of single-sex secondary schools on student achievement and attitudes." Journal of Educational Psychology, 78(5): 381-95.

Lee, V. E. and M. E. Lockheed (1990). "The effects of single-sex schooling on achievement and attitudes in nigeria." Comparative Education Review, 34(2): 209-31.

Lefgren, L. (2004). "Educational peer effects and the Chicago public schools." Journal of Urban Economics, 56(2): 169-91.
Link, C. R. and J. G. Mulligan (1991). "Classmates' effects on black student achievement in public school classrooms." Economics of Education Review, 10(4): 297-310.
Manski, C. F. (1993). "Identification of endogenous social effects: The reflection problem." Review of Economic Studies, 116(2): 607-54.
Manski, C. F. (2000). "Economic analysis of social interactions." Journal of Economic Perspectives, 13(3): 115-36.
McEwan, P. (2003). "Peer effects on student achievement: Evidence from Chile." Economics of Education Review, 22(2): 131-41.
Neidell, M. and J. Waldfogel (2010). "Cognitive and noncognitive peer effects in early education." Review of Economics and Statistics, 92(3): 562-76.
Niederle, M. and L. Vesterlund (2007). "Do women shy away from competition? Do men compete too much?" Quarterly Journal of Economics, 122(3): 1067-101.
Niederle, M. and L. Vesterlund (2010). "Explaining the gender gap in math test scores: The role of competition." Journal of Economic Perspectives, 24(2): 129-44.
Rivkin, S. G., E. A. Hanushek and J. F. Kain (2005). "Teachers, schools, and academic achievement." Econometrica, 73(2): 417-58.
Rockoff, J. (2004). "The impact of individual teachers on student achievement: Evidence from panel data." American Economic Review, 94(2): 247-52.
Rockoff, J. (2009). "Field experiments in class size from the early twentieth century." Journal of Economic Perspectives, 23(4): 211-30.
Rotter, J. (1966). "Generalized expectancies for internal versus external control of reinforcement." Psychological Monographs, 80: 1-28.
Schneeweis, N. and M. Zweimüller (2009). "Girls, girls, girls: Gender composition and female school choice." University of Linz Economics Working Papers 05.
Schütz, G., H. Ursprung and L. Wößmann (2008). "Education policy and equality of opportunity." Kyklos, 61(2): 279-308.
Smith, P. B., S. Dugan and F. Trompenaars (1997). "Locus of control and affectivity by gender and occupational status: A 14 Nation Study." Sex Roles, 36(1-2): 51-77.
Steele, C. M. (1997). "A threat in the air. How stereotypes shape intellectual identity and performance." American Psychologist, 52(6): 613-29.
Sutter, M. and D. Rützler (2010). "Gender differences in competition emerge early in life." University of Innsbruck, Working Papers in Economics and Statistics 14.
Trautwein, U., O. Lüdtke, O. Köller and J. Baumert (2006a). "Self-esteem, academic self-concept, and achievement: How the learning

environment moderates the dynamics of self-concept." Journal of Personality and Social Psychology, 90(2): 334-49.
Trautwein, U., O. Lüdtke, H. W. Marsh, O. Köller and J. Baumert (2006b). "Tracking, grading, and student motivation: Using group composition and status to predict self-concept and interest in ninth-grade mathematics." Journal of Educational Psychology, 98(4): 788-806.
Trautwein, U., O. Lüdtke, H. W. Marsh and G. Nagy (2009). "Within-school social comparison: How students perceive the standing of their class predicts academic self-concept." Journal of Educational Psychology, 101(4): 853-66.
Wigfield, A. and M. Karpathian (1991). "Who am I and what can I do? Children's self-concepts and motivation in achievement situations." Educational Psychologist, 26: 233-62.

THURGAU INSTITUTE
OF ECONOMICS
at the University of Konstanz

Hauptstr. 90
CH-8280 Kreuzlingen 2

Telefon: +41 (0)71 677 05 10
Telefax: +41 (0)71 677 05 11

info@twi-kreuzlingen.ch
www.twi-kreuzlingen.ch