Mechanisms and Impacts of Gender Peer Effects at School

By Victor Lavy and Analía Schlosser*

Abstract

We present in this paper evidence about the effects and mechanisms of gender peer effects in elementary, middle, and high schools. For identification, we rely on idiosyncratic variations in gender composition across adjacent cohorts within the same schools. We find that an increase in the proportion of girls improves boys and girls' cognitive outcomes. These academic gains are mediated through lower levels of classroom disruption and violence, improved inter-student and student-teacher relationships, and lessened teachers' fatigue. We find no effect on individual behavior, which suggests that the positive effects of girls on classroom environment are mostly due to compositional change.

---

* Lavy: Department of Economics, Hebrew University of Jerusalem, Mount Scopus, Jerusalem 91905, Israel, Royal Holloway, University of London, Bureau for Research and Economic Analysis of Development (BREAD), Center for Economics and Policy Research (CEPR), and National Bureau of Economic Research (NBER) (e-mail: msvictor@mscc.huji.ac.il); Schlosser: The Eitan Berglas School of Economics, Tel Aviv University, POB 39040, Ramat Aviv, Tel Aviv, 69978, Israel, and Institute for the Study of Labor (IZA) (e-mail: analias@post.tau.ac.il). We thank the Ministry of Education for assisting with the data, Rachel Berner, Issi Romem, and Yannay Spitzer for outstanding research assistance, and comments from Kirk Doran, Alfronso Flores-Lagunes, Alan Krueger, Cecilia Rouse, and seminar participants at LSE, the Institute of Education at the University of London, Tel Aviv University, the Hebrew University, Princeton, MIT, Yale, the 2007 SOLE meetings, the Bergen-CEPR 2007 conference, the LEaF conference, and the IZA/SOLE transatlantic meeting. Schlosser thanks the Industrial Relations Section and the Pinhas Sapir Center for research support.

I. Introduction

The question of whether classroom gender-composition matters for student learning has long been of concern to social scientists, educators, and policymakers. The general view is that social interactions between genders at school often play an important role in academic achievement and career choices. However, little scientific evidence supports these beliefs and not much is known about the mechanisms of these peer effects. Such evidence is more relevant now given the concern about gender imbalances in co-educational schools due to the recent increase in the number of single sex classes and schools.[1] While much attention has been given to the comparison of students outcomes in single sex and co-educational classes, a recent report by the American Association of University Women indicates that an overlooked consequence in the creation of single-sex classes is the disruption of the sex ratio in co-educational classes from which single-sex classes are drawn (Susana Morse 1998). [2] This phenomenon has already been noticed in the United Kingdom where a higher demand for single sex-schools for girls relative to boys has resulted in highly imbalanced sex ratios in some co-educational public schools.[3] In Inner London, for example, a higher ratio of girls in single-sex schools is reflected in co-educational schools, where 59 percent of the students are boys. Understanding the effects of classroom gender composition is therefore important to assess the consequences of imbalanced

[1] In the US, for example, there is an increasing trend in the number of single-sex schools as a response to the new Title IX single-sex regulations released in October 2006 which give communities more flexibility in offering single-sex classes and permit school districts to provide single-sex schools. For more details, see http://www.ed.gov/news/pressreleases/2006/10/10242006.html (accessed November 14, 2010)

[2] See the National Association for Single Sex Public Education website: http://www.singlesexschools.org (accessed November 14, 2010) and Patricia Campbell and Jo Sanders (2002) for a discussion of the pros and cons of single-sex schooling.

[3] A recent article in The Guardian (April 10, 2007), discusses the effect of single-sex schools on the gender imbalance in public schools in the UK and explains that the higher proportion of girls in single-sex schools relative to boys reflects the desire of parents to send their daughters to single-sex schools, but not their sons. See: http://education.guardian.co.uk/egweekly/story/0,,2053138,00.html (accessed November 14, 2010)

sex ratios in co-educational public schools and to determine an optimal grouping of students into classrooms and an efficient allocation of resources within and across schools.

This paper examines the extent of gender peer effects in the educational production function. The first part of the paper investigates how classroom gender composition affects the scholastic achievements of boys and girls in different stages of the schooling cycle. As outcomes in elementary school and in middle school, we use test scores in English, Hebrew, math, and science for 5th and 8th grades. For high school, we use several measures of students' performance in the matriculation exams.

The second part of the paper explores mechanisms by which gender peer composition affects academic outcomes. Our study appears to be the first to uncover the "black box" of peer effects, particularly those that derive from classroom gender composition. We focus on the following mechanisms: classroom disruption and violence, inter-student interactions, student-teacher relationships, and teachers' sense of "fatigue" or "burnout" with their job. This form of externalities of the presence of girls in the classroom is a reflection of the *congestion effect* in the education production model proposed by Edward Lazear (2001). However, the peer effects of girls can also results from changes in the probability that a student misbehaves, which in Lazear's model is assumed to be fixed. We are able to disentangle these two channels of the peer effect by distinguishing between the effects generated by changes in classroom gender composition and those caused by changes in the behavior of students. This analysis is based on contrasting students' views about their classroom environment with students' reports on their own behavior.

To control for unobserved characteristics of schools and students that might be correlated with peer gender composition and that may also affect students' outcomes, we rely on idiosyncratic variations in the proportion of female students across adjacent cohorts within the same school. By using multiple cohorts and conditioning on school fixed effects and school-specific time trends, we are able to control for unobserved factors that might confound the gender peer effect in schools. Using Monte Carlo simulations, we show that this within-school variation in the proportion of female students resembles the variation that would be generated by a random process. We further demonstrate that within-school variation in the proportion of girls is not related to within-school variation in student background characteristics providing additional evidence supporting the validity of our identification strategy. We also show that mobility rates of students across schools are very low in Israel, which makes the identification strategy particularly attractive in this context, especially as we also demonstrate that the proportion of girls in a student's grade does not affect the likelihood of a student's mobility.

In studying the mechanisms of gender peer effects we are able to exploit an additional identification strategy based on longitudinal data that allow us to observe the same students in two different school environments, elementary and middle school. In this case, we generate student fixed effects estimates for the impacts of variation in the proportion of female peers that result from student's transition from elementary to middle schools.

The empirical evidence on gender peer effects in schools is based primarily on studies that contrast outcomes for students, usually for girls, in single-sex and co-educational classes. The US Department of Education (2005) and Susan Morse (1998) review such studies in elementary and high schools, and Irene Harwarth, Mindi Maline, and Elizabeth Debra (1997) includes a review of studies in colleges. The evidence is mixed; some studies suggest no differences between single-sex and co-educational schooling while others find that single-sex schooling may be beneficial. Evidence favoring co-educational schooling is much more limited. Nevertheless, it is difficult to interpret these findings since most of the studies do not account for the non-random selection of students into single-sex and co-educational schools and unobserved potentially confounding differences between these two types of institutions that may generate *correlated effects* (Charles Manski 1993) and be confounded with peer effects.

Some recent studies use experimental or quasi-experimental research designs to separate the *social effects* in the classroom from the *correlated effects* (see, e.g., Bruce Sacerdote 2001; David J. Zimmerman 2003; Joshua Angrist and Kevin Lang 2004; Peter Arcidiacono and Sean Nicholson, 2005; Eric Hanushek, John F. Kain, Jacob Markman, and Steven Rivkin 2003; Eric Gould, Victor Lavy, and Daniele Paserman 2009; Caroline M. Hoxby and Gretchen Weingarth 2005; and Ammermueller and Pischke, 2009). However, only a few studies focus on gender peer effects.

The only exception is Caroline M. Hoxby (2000) who estimates gender and race peer effects in Texas elementary public schools and finds that boys and girls have higher test scores when classrooms have more female students.[4] Our study complements and extends her work in at least four important dimensions. First, we are able to study the effects of classroom gender composition at all three levels of education (elementary, middle schools, and high schools). This is of paramount importance since gender interactions change considerably by age and therefore, it is not clear that the evidence on gender peer effects in elementary schools is relevant for higher levels of education. Second, we use a national sample of all public schools in an educational system where students' mobility between schools is extremely low and where private schools are virtually inexistent. We are therefore able to minimize the problems of students' endogenous mobility according to classroom gender composition; an issue of concern in previous work on this topic. Third, using unique data on non-cognitive outcomes, we go a

---

[4] A second study that examined gender peer effects is Diane Whitmore (2005) who finds mixed results for the effects of the proportion of female students using gender variation across classrooms generated by Tennessee's Project STAR.

step further in the estimation of peer effects and examine the mechanisms through which peers affect students' scholastic achievement. In the examination of mechanisms, we also introduce an alternative empirical strategy based on a student fixed effects approach, where we estimate the impacts of changes in classroom gender composition generated by students' transition from elementary to middle school. In this case, we do not only look at gender peer effects within a fixed school environment but we are also able to examine the impacts of changes in peers' gender composition within the same student.

The results we present in the paper show that the proportion of girls in a class has a positive and significant effect on the academic achievements of girls and of boys. The sizes of the estimated effects are similar for both genders suggesting that a change in classroom gender composition could be close to a zero sum gain as boys benefit from being with more girls but girls benefit from having fewer boys in class. Nevertheless, we find sharp heterogeneous effects by students' socioeconomic background that show larger benefits for students with low parental education and for new immigrants.

All estimated effects are significantly different from the results of falsification tests that use placebo treatments, which show no effect at all. These falsification tests are based on replacing the treatment variable with the proportion of females in the previous or the subsequent cohort in the same school. The lack of any discernable effects when the placebo treatments are used, suggests that the estimated treatment effects are not spuriously picking up any short-term effects of unobserved confounders at the school level.

An examination of the underlying mechanisms of the gender peer effects shows that a higher proportion of girls in the classroom lowers the level of classroom disruption and violence, and improves inter-student and teacher-student relationships. The improvement in classroom environment is also reflected in lower levels of teachers' fatigue and feelings of burnout, On the other hand, the estimates of the effect of the proportion of girls on students' (self-reported) violent behavior, disciplinary problems, and study effort show no systematic or significant relationship. The sharp contrast between these two sets of results, suggests that much of the improvement in the classroom environment associated with a higher proportion of girls is due to a change in classroom gender composition and not to changes in individual student behavior.

The rest of the paper is organized as follows: Section II describes the identification strategy. Section III discusses the data, the construction of the analysis samples, and presents various pieces of evidence that assess the validity of our identification strategy. Section IV reports the school fixed effects estimates of gender peer effects on elementary, middle, and high school students' achievement, while section V presents evidence on the possible mechanisms driving the positive female peer effects on students' achievement. Section VI shows results suggesting that a change in classroom gender composition and not behavioral changes among students is the driving force behind the estimated gender peer effects on classroom environment. Section VII concludes.

II. Empirical Strategy

A. Identification of Gender Peer Effects

The effect of classroom gender composition on students' outcomes is usually confounded by the effects of unobserved correlated factors. Such correlations could result if self-selection and sorting of students across schools are affected by school gender composition or if school gender composition is correlated with other characteristics of the school that may affect students' outcomes. One possible method to account for both sources of confounding factors in the estimation of peer effects is to rely on within school variations in the proportion of female students across adjacent cohorts.[5] Based on this approach, we examine whether cohort-to-cohort changes in male and female outcomes within the same grade and school are systematically associated with cohort-to-cohort changes in the proportion of female students. The basic idea is to compare the outcomes of students from adjacent cohorts who have similar characteristics and face the same school environment, except for the fact that one cohort has more female students than the other due to purely random factors.

While implementing this methodology, we use the proportion of female students measured at the grade and not at the classroom level, because the latter might be endogenous as parents and school authorities may have some discretion in placing students in different classes within a grade. This is not a very restrictive compromise because within a given school the proportion of female students in a grade is highly correlated with the proportion of female students in a class.[6]

Using repeated cross-sectional data, we estimate the following reduced-form equation separately for boys and girls and for separate samples of elementary, middle, and high school students to explore how gender peer effects evolve through the different schooling stages:

$$(1)\ y_{igst} = \alpha_g + \beta_s + \gamma_t + x_{igst}^{'}\lambda_1 + S_{gst}^{'}\lambda_2 + \pi P_{gst} + \varepsilon_{igst}$$

where *i* denotes individuals, *g* denotes grades, *s* denotes schools, and *t* denotes time. y_{igst} is an achievement measure for a male/female student *i* in grade *g*, school *s*, and year *t*; α_g is a grade effect, β_s is a school effect, γ_t is a time effect, x_{igst} is a vector of student's covariates that

---

[5] A similar identification strategy was recently applied by Hoxby (2000) to estimate gender and race peer effects in elementary schools in Texas. Other studies that rely on within school variation in peer composition are Angrist and Lang (2004); Gould, Lavy, and Paserman (2009); and Ammermuller and Pischke (2009).

[6] The correlation between the proportion of female students in the grade and the proportion of female students in the class is 0.67 for elementary schools. The correlation for middle schools and high schools is 0.56 and 0.55 respectively. Nevertheless, we think that at higher levels of education, the proportion of female students in the grade (and not in the class) is a more relevant measure of treatment since students spend a lower proportion of the school day in their homeroom class.

includes mother's and father's years of schooling, number of siblings, immigration status, and ethnic origin, and indicators for missing values in these covariates, S_{gst} is a vector of characteristics of a grade *g* in school *s* and time *t* and includes a quadratic function of enrollment and a set of variables for the average characteristics of the students in the grade; P_{gst} is the proportion of female students in grade *g* (which we refer to as the proportion female from here on), school *s*, and year *t*, and ε_{igst} is the error term, which is composed of a school-specific random element that allows for any type of correlation within observations of the same school across time and an individual random element. The coefficient of interest is π, which captures the effects of having more female peers on student achievement.

For the estimates in equation (1) to have a causal interpretation, the unobserved determinant of achievement must be uncorrelated with the treatment variable. Including school fixed effects controls for the most obvious potential confounding factor – the endogenous sorting of students across schools. However, one may be concerned that there are time-varying unobserved factors that are also correlated with changes in the proportion of female students. To address this concern, we add to equation (1) a full set of school-specific linear time trends (δ_s). In this case, identification is achieved from the deviation in the proportion of female students from its school long-term trend and is estimated by the following equation:[7]

$$(2)\ y_{igst} = \alpha_g + \beta_s + \delta_s year_{st} + \gamma_t + x'_{igst}\lambda_1 + S'_{gst}\lambda_2 + \pi P_{gst} + \varepsilon_{igst}$$

B. *Identification of Mechanisms*

[7] Equation (2) is estimated for high school outcomes because we have a longer panel and also because secular trends in school gender composition are more likely to exist since there is school choice at this level of education.

The parameter π in equations (1) and (2) measures gender peer effects that could be enacted through various channels. This could include effects through changes in the classroom climate, in the quality of interactions among students and between students and teachers, in the level of motivation and self-confidence of students; through modifications in students' effort and study habits; and also through responses of teachers in terms of their effort, attitudes towards the class, and teaching methods. To assess the importance of each of these mechanisms, we estimate models identical to model (1) where the dependent variables are constructed from students' responses to a school questionnaire about classroom environment, study efforts, and their own behavior, as well as from teachers' reports about their sense of fatigue and work satisfaction.

It is important to note that the mechanisms through which gender peer effects may operate can simply reflect a change in classroom gender composition but can also reflect changes in the individual behavior of students. For example, a higher proportion of girls in the classroom can improve the classroom climate by lowering the incidence of disruptions simply because there are fewer boys, who tend to be more disruptive than girls. In addition, having more girls in a class may affect students' individual behavior. A violent boy may be more tranquil and less disruptive due to a more relaxed atmosphere that girls may create or because teachers may be more patient with more girls in the class. These behavioral changes impact the class environment in addition to the compositional effect described above.

We propose to disentangle these two alternative explanations by using two different types of questions in the student questionnaire. In one set of questions, students are asked about their views regarding general aspects of their classroom (for example, the level of violence). The effect of the gender mix on these measures captures the overall gender peer effect (due to compositional changes and changes in students' behavior). In another set of questions, students are asked about their own behavior (for example, if they were involved in a violent interaction during the current year). We interpret the effect of classroom gender composition on measures of students' own, self-reported behavior as indications of changes in individual behavior. More details about these questions are provided in the next section.

III. Data

The empirical analysis is based on three samples that include elementary, middle, and high school students, respectively. All three samples include only schools that have mixed-gender classes because the identification strategy is based on within school variation in the proportion of female students. This condition is met in all Jewish secular elementary, middle, and high public schools and in about 50 percent of the Jewish religious elementary public schools. A small number of religious schools have mixed-sex classes at the middle and high school level but since this sample is very selective, we prefer not to include them in the analysis. Below we describe the three samples.

A. The High School Data

We use administrative records collected by the Israel Ministry of Education for eight consecutive cohorts (from 1993 to 2000) of 10th grade students. The data are based on annual reports submitted by school authorities to the Ministry of Education at the beginning of the school year. Each record contains an individual identifier, a school and class identifier, and detailed demographic information on the student: gender, parental education, number of siblings, year of immigration (where relevant), and ethnicity. We use 10th grade to define the base population because it is the first year of high school and the last year of compulsory schooling. The measure of treatment in high school in terms of the proportion of female peers is also based on 10th-grade enrollment because any later change in this rate is endogenous. The sample is restricted to students in non-special education classes in secular schools that have a matriculation track. As a further restriction, we drop all schools that experienced a change in enrollment of 80 percent or more between 2 consecutive years of the analyzed period to avoid changes in school gender composition that might have originated from structural changes in the school. In addition, we drop schools that have an annual enrollment lower than 10 students.

Israeli high school students are enrolled either in an academic track leading to a matriculation certificate (*Bagrut* in Hebrew) or in an alternative track leading only to a high school diploma.[8] The *Bagrut* is completed by passing a series of national exams in core and elective subjects taken by the students between 10th and 12th grade.[9] Students choose to be tested at various levels of proficiency, with each test awarding from one to five credit units per subject, depending on difficulty. Some subjects are mandatory, and for many the most basic level is three credit units. Advanced level subjects are those subjects taken at four or five credit units. A minimum of 20 credit units is required to qualify for a matriculation certificate. We link the students' datasets with administrative records that include the results (test scores) of these matriculation exams.

We focus on the following matriculation outcomes that are available for all the years: the average score in the matriculation exams, matriculation status (=1 if awarded with the matriculation diploma and 0 otherwise), number of credit units, number of advanced level

---

[8] The matriculation certificate is a prerequisite for university admission and receiving it is one of the most economically important educational milestones. Similar high school matriculation exams are found in many countries and in some states in the United States. Examples include the French Baccalaureate, the German Certificate of Maturity, the Italian Diploma di Maturità, and the New York State Regents examinations.

[9] The matriculation tests are national exams written and scored by an independent agency. Therefore the average score of students is not affected by the within school distribution of test scores. The same argument applies for the test score data used in the analysis for elementary and middle schools and described below.

subjects in science, and matriculation status that meets university entrance requirements (at least 4 credits in English and another subject at a level of 4 or 5 credits).[10]

B. The Middle and Elementary School Data

Data for elementary and middle schools are based on the GEMS (Growth and Effectiveness Measures for Schools - *Meizav* in Hebrew) datasets for the years 2002-2005. The GEMS includes a series of tests and questionnaires administered by the Division of Evaluation and Measurement of the Ministry of Education.[11] The GEMS is administered at the midterm of each school year to a representative 1-in-2 sample of all elementary and middle schools in Israel, so that each school participates in GEMS once every two years.

The GEMS student data include test scores of 5th and 8th graders in math, science, Hebrew, and English, as well as the responses of 5th through 9th grade students to questionnaires. In principle, all students except those in special education classes are tested and required to complete the questionnaire. The proportion of students who are tested is above 90 percent, and the rate of questionnaire completion is roughly 91 percent. The raw test scores used a 1-to-100 scale that we transform into z-scores to facilitate interpretation of the results.

The GEMS student questionnaire addresses various aspects of the school and learning environment. We select a section that focuses on the classroom climate and student behavior. In this section, students are asked to rate in a 6-point scale ranging from 1 (strongly agree) to 6 (strongly disagree) the extent to which they agree with a series of statements. We also examine a

---

[10] Roughly, 10 percent of the students in the sample did not take any of the matriculation exams. These students get zero values in the average score. None of the other four matriculation outcomes require such imputation since the zero values that these students get for these outcomes is a real and not an imputed measure of achievement.

[11] The GEMS are not administered for school accountability purposes and only aggregated results at the district level are published. For more information on the GEMS see the Division of Evaluation and Measurement website (in Hebrew): http://cms.education.gov.il/educationcms/units/rama/odotrama/odot.htm (accessed November 14, 2010).

set of items in the questionnaire where students report the amount of time allocated to homework in math, Hebrew, English, and science and technology.

The student questionnaire data and test scores for the years 2002-2005 were linked to student administrative records collected by the Israel Ministry of Education (identical in structure to the data used for high school students). The administrative records include student demographics and are used to construct peer gender composition and all measures of students' background characteristics. Using the linked datasets, we built a panel for elementary schools and a panel for middle schools. As we did for the high school sample, we drop any schools with an annual enrollment lower than 10 students from the panel.

The elementary school panel includes data from 5th- and 6th-grade student questionnaires and 5th-grade student test scores for the years 2002-2005. The sample is restricted to Jewish public schools that have mixed-gender classes. There are 997 elementary schools (808 secular and 189 religious) with test score data and 1,010 elementary schools (808 secular and 202 religious) with student-questionnaire data. Since every school is sampled once in two years, we have two observations of the same school and grade for more than 90 percent of the schools.

The middle school panel includes student questionnaires for 7th through 9th grades and 8th-grade student test scores for the years 2002-2005. The sample is restricted to secular schools, since there are only a few religious middle schools with mixed-gender classes. There are 395 secular schools in the sample, of which 85 percent appear in two years.

As we have multiple grades for each school in the student's questionnaire data, we pool all grades and years and exploit within school variation in the proportion of female students across grades and years to gain more variability in this variable. We therefore have four observations of the same school for elementary schools (5th and 6th grade for two years) and six observations of the same school for middle schools (7th, 8th, and 9th grade for two years). The analysis on student test scores for elementary and middle schools has more limited power because only one grade was tested, leaving us with only two observations per school.

The GEMS also includes interviews with all teachers and the school principal. The teacher survey included mainly questions about resources for instruction and training, but it also included three questions about teaching fatigue ("burnout"), workload, and overall work satisfaction. We use teachers' responses to these items to explore another mechanism of the gender peer effect: namely, whether the proportion of girls in the classroom affects teachers' fatigue and work satisfaction, which are likely to be correlated with teachers' unobserved productivity.

C. Evidence on the Validity of the Identification Strategy

Our key identifying assumption postulates that changes in the proportion female within a school are uncorrelated with changes in unobserved factors that could affect students' outcomes. We assess here, from different angles, the feasibility of this assumption. We first examine the source of the within school variation in the proportion of female students. We argue that in this regard, idiosyncratic fluctuations in the gender composition of incoming cohorts in a school generate this variation.[12] That is, while the proportion of female students in a school is relatively stable over time, there are year-to-year deviations for each incoming cohort that are mostly generated by natural fluctuations in the number of boys and girls of a particular birth cohort who

[12] In Victor Lavy and Analia Schlosser (2007) we show that, as expected, the variation in the proportion female is larger in small schools, but is also evident in medium and large schools. In addition, there are schools with a significant amount of variation located in small and large towns as well as in the main metropolitan areas.

live in a school catchment area.[13] These differences in the gender composition across incoming cohorts persist through their progression to higher grades in the same school.

To illustrate this point, we show in Figure 1a that the within school variation in the proportion female over the years 2002-2006 is virtually identical in 1st and 5th grade and it is similar to the variation in the proportion of girls aged 6 that lives within a residential area over the same period.[14] We also performed Monte Carlo Simulations for the elementary, middle, and high school sample to assess whether the observed within school variation in the proportion female resembles the variation that would result if the gender composition of each cohort was randomly generated.[15] The result of one such simulation, plotted in Figure 1b, clearly shows that the actual within school variation in the proportion female in elementary schools is virtually identical to the simulated variation. Based on these simulations, we also computed an empirical confidence interval for the standard deviation in the proportion female, finding that 89% of the schools in our sample had a standard deviation that fell within the empirical 90% confidence interval, which is close to our expectations.[16]

---

[13] Figures 1 and 2 of the Web Appendix show the school average and standard deviation in the proportion female by grade. We also report in Web Appendix Table 1 the standard deviation in the proportion female for each grade and the extent left after removing school fixed effects. In elementary and middle schools, about 83%-90% of the overall standard deviation in the proportion female is within schools since every school that has mixed-gender classes is expected to have an equal proportion of male and female students, so that between school variations are relatively small. At the high school level, the variation in the proportion female is larger between than within schools since there is some sorting by gender of students across schools and because average school enrollment is higher at this level of education.

[14] Residential areas are defined by zip codes. The variation in the proportion female is smaller within residential areas than within schools because the average cohort size is larger in a residential area than in a school.

[15] For each school, we randomly generate the gender of the students in each cohort using a binomial distribution function with p equals to the average proportion of females in the school across all years. We then compute the within school standard deviation of the proportion female and repeat this process 1,000 times to obtain an empirical 90 percent confidence interval for the standard deviation for each school. For the high school sample, we compute within school standard deviations using residuals from a regression of the proportion female on school fixed effects and school specific time trends.

[16]The results for the middle school and high school samples are virtually identical and are available from the authors upon request. We further re-estimate all models by restricting the sample to schools where the standard deviation falls within its confidence interval and we obtain virtually identical results to those obtained based on the full sample and reported below.

[Figures 1a and 1b]

Even if the fluctuations in the proportion female within a school resemble a random process, these variations could be correlated with additional cohort-to-cohort changes that might affect student outcomes. To assess this possibility, we check whether changes in the proportion of girls within a school are associated with changes in student background characteristics such as parental education, family size, ethnicity, and the proportion of new immigrants. Table 1 provides evidence on these balancing tests by reporting the estimated coefficients from within school regressions (by including school fixed effects) of various student characteristics on the proportion of female students in elementary, middle, and high school. OLS estimates are also reported, as a benchmark for comparison.[17]

[Table 1]

In the elementary school sample, the proportion of female students in a grade is not related to most of the observable student characteristics, both in the OLS and the within school (fixed-effects) regressions. The only exception seems to be a negative association between the proportion of female students and the proportion of students from Asian/African origin. However, this association is largely reduced and becomes insignificant when adding school fixed effects. In the middle school sample, the OLS estimates suggest that grades with a higher proportion of female students have a lower proportion of new immigrants and a higher proportion of students from Asian/African ethnicity. These negative correlations, however, are virtually zero and insignificant in the within school regressions.

At the high school level, the OLS estimates show some associations between school gender composition and student background characteristics. However, these correlations are

[17] We also perform similar balancing tests in sub samples stratified by gender and do not find any association between within school changes in the proportion of girls and changes in the background characteristics of boys or girls.

largely reduced and became insignificant in the within school regressions. The addition of school-specific linear time trends eliminates all associations. For example, the coefficient on father's years of schooling is 0.606 (s.e.=0.648) in the OLS regression. It drops to 0.517 (s.e.=0.445) in the within school regression and it is further reduced to -0.097 (s.e.=0.414) when adding school specific linear time trends. Overall, the results for elementary schools, middle schools, and high schools show that cohort-to-cohort changes in the proportion of female students within a school appear to be uncorrelated with other changes in student background characteristics.[18]

We also check whether changes in the proportion female are associated with changes in school enrollment. As reported in the last row of Table 1, there are some imbalances according to this variable but they have opposite signs in the different sub-samples. For example, we observe a positive association between the proportion female and enrollment for elementary schools that becomes marginally significant only when adding school fixed effects. On the other hand, there is no association between the two variables at the middle school level, while there is a negative association at the high school level that becomes marginally significant only when adding school fixed effects and school specific time trends. Given the inconsistency across samples and specifications, we interpret these associations as spurious. In any case, in all outcome regressions we control for a quadratic function of enrollment.[19]

Even if cohort-to-cohort variations in the proportion female could be purely idiosyncratic within a school, one could still be concerned that students might respond to these

---

[18] There could of course be a systematic correlation between students' unobservables and the proportion of female students. We cannot entirely rule out this possibility, even though the lack of a correlation in the observables hints that the presence of a strong correlation in the unobservables is very unlikely, especially if these unobservables are correlated with the observed covariates.

[19] It is also worth noting in this regard, that the quadratic function of enrollment does not have a significant effect in any of the outcome regressions. This fact further reduces the concern of possible biases.

unpredicted shocks to cohort gender composition. The lack of school choice at the elementary and middle school level and the very limited scope of private schooling in Israel, significantly diminishes this concern. In high schools, such selection could potentially occur, but it is very unlikely because while parents may know the average gender composition of a school, it will be difficult for them to predict in advance the gender composition of a cohort that enters the school in a particular year. Nevertheless, they could still leave a school after they are exposed to this information, in all likelihood after the beginning of the school year. We therefore address this concern by checking whether the likelihood that a student leaves a school (by moving to another school or dropping out) is associated with the proportion of female students in his/her initial grade. We focus on three key enrollment decisions, entry to the first grade of elementary, middle, and high school, and construct a dummy variable that equals one if the student left the school in the following year.[20] Using this indicator as a dependent variable, we estimate models similar to (1) and (2) to assess the effects of the proportion of female students in the grade on the likelihood that a 1st, 7th or 10th grade student leaves his/her initial school.[21]

Table 2 reports the regression results along with the outcome means. The first fact to note is that the rate of students' mobility is relatively low. Roughly, 8 percent of the students left their school at the transition between 1st and 2nd grade and the rates for 7th and 10th grade are 5 and 8 percent respectively (see first row of the table). The low mobility rates in comparison, to

---

[20] To avoid classifying as school movements or drop-outs those cases that arise from structural school changes (closures, merges, etc.) or from data collection problems, we follow Eric A. Hanusheck, John F. Kain, and Steven G. Rivkin (2004a) and exclude from school leavers those cases where the student moved to a school attended by more than 30 percent of the students of his/her former grade. We further excluded from school leavers those cases where 100 percent of the students in the grade left the school. Less than half percent of the sample is affected by these two adjustments.

[21] For the model at the high school level, we are able to use the exact same cohorts that are used to produce the results reported in the next section. For middle schools, we have only three years of data with student IDs that were traceable over time so the model includes the 2001 and 2002 7th grade cohorts and their follow-up in 2002 and 2003, respectively. At the elementary school level we have only two years of data with student IDs that were traceable over time (2002 and 2003) leaving us with only one cohort for the follow-up of 1st graders. Therefore, the analysis for elementary schools does not include school fixed effects.

the US for example, make the implementation of an identification strategy based on within school variation in the proportion of female students especially appealing in the Israeli context.[22] The estimates of the effects of the proportion female on the likelihood of leaving the initial school are reported in second row of the table. All estimates are small, insignificant, and sometimes have opposite signs across different grades. Overall, these results suggest that the likelihood that a student leaves his/her initial school is unrelated to the proportion of female students in his/her cohort.

[Table 2]

IV. Results

A. Effects on High School Students' Achievement

Table 3 reports the effects of the proportion of female peers on high school achievements. The sample includes 264 high schools and 404,929 students from eight cohorts. The proportion of female students is roughly 50 percent in all the cohorts and it has no apparent time trend. Columns 1 and 4 present the outcome means for girls and boys, respectively. Female students consistently outperform males in almost all matriculation outcomes except in the number of advanced level subjects in science.

Columns 2-3 and 5-6 report the effects of the proportion female on girls and on boys' matriculation outcomes respectively. The estimates presented are based on two different specifications. Columns 2 and 5 report estimates when year dummies, school fixed effects, school specific time trends, school enrollment, and individual's and cohort mean characteristics are included as controls. In order to assess how sensitive these estimates are to the control of

---

[22] A US national study reports that 40 percent of third graders have changed schools at least once since 1st grade (US General Accounting Office, 1994). Hanusheck, Kain, and Rivkin (2004a) report an annual rate of student mobility of 24% in Texas elementary schools. Similar annual rates are reported for Ohio by Virginia Rhodes (2005) and for Florida in personal conversation with David Filgio.

individual and cohort characteristics we report estimates in columns 3 and 6 based on a specification that excludes them from the regression.[23]

[Table 3]

Focusing on the estimates from the complete specification (columns 2 and 5) we see that both females and males tend to perform better in each of the five high school outcomes when they are in classes with a higher proportion female. Three of the five estimates for girls are significantly different from zero at the 5 percent level and the other two are significant at the 10 percent level. The effect on boys is also positive for all five outcomes; three of them are precisely measured. Noteworthy is the similarity of the estimates for boys and girls, for example, the effect on credit units is 1.5 for girls and 1.4 for boys.[24]

Column 3 and 6 present the estimates when we omit the student and cohort characteristics as controls. The effect sizes are nearly identical in comparison to those reported in columns 2 and 5 while the estimated standard errors are smaller in the more inclusive specifications. This pattern is replicated in other estimates that we present in the paper. The robustness of the estimates with respect to these controls is a result of the well-balanced characteristics with respect to the proportion of girls in the cohort once we control for school fixed effect and school specific linear time trends. Adding the covariates as controls improves the precision of the estimates in the same way that regression adjusting increases precision in an experimental setting.[25]

---

[23] We also estimate models similar to those presented in columns 2-3 and 5-6 based on aggregate data at the school/year/gender level weighted by cell size. These results (not reported here to save space) are almost identical to the results using micro data.

[24] We fail to reject the null hypothesis of equality of the boys and girls' estimates for each of the five matriculation outcomes. The hypothesis tests are based on the estimation of seemingly unrelated regressions to account for the correlation between the estimates for boys and girls.

[25] We also estimate three alternative versions of the full model reported in columns 2 and 5 where we use different controls for the average background characteristics of the cohort. In one model, we control separately for the average characteristics of girls and boys. In two additional specifications, we alternate and control for the average

The above estimates imply effects of moderate size. For example, a 10 percentage point increase in the proportion of female peers increases the probability of matriculation by almost one percentage point among girls, and by half a percentage point among boys. To put this in perspective, assuming that the gender peer effects are linear, the estimates suggest that an all-female class would increase the matriculation rate of girls by about 10 percentage points. Though in absolute terms it is a moderate impact, it is not so in comparison to the gains obtained from recent educational interventions aimed at raising the matriculation rate. For example, a 20 percentage point increase in the proportion of female peers would lead to an increase in the probability of matriculation of 1-2 percentage points. This effect is half of the effect size estimated by Victor Lavy and Analia Schlosser (2005) for a remedial education program that provided additional instructional hours to high school students and a quarter of the size of that estimated by Angrist and Lavy (2009) for a program that provided large conditional monetary bonuses to high school students.

Another example that highlights the relative size of the effect uses the estimates of the average score for females (6.314) and for males (7.918), which imply that a 20 percentage point increase in the proportion of female peers, increases average scores of girls by 1.3 points and average scores of boys by 1.6 points. These gains imply an approximate increase of 4-5 percent of a standard deviation in the students' test score distribution. An all-female class would raise the score of girls by 0.20-0.25 of a standard deviation, similar to the effect of reducing class size by 33 percent (Joshua Angrist and Victor Lavy 1999).

---

characteristics of boys or girls in the cohort. All estimates of these three alternative models are virtually identical to those obtained when controlling for the average characteristics of the cohort.

B. ***Falsification Tests***

Columns 7-10 of Table 3, present the falsification tests based on placebo measures of treatment, namely when the proportion of female students in the younger cohort (t-1) or the older cohort (t+1) replaces the true treatment measure. [26] The results based on the t-1 or t+1 measure of treatment show no effect on any of the outcomes, for boys and for girls. All estimates are small, have inconsistent signs, and are insignificant. For example, when using the proportion of girls of the t+1 cohort (columns 8 and 10) the estimates of the matriculation rate are 0.028 (s.e.=0.046) for girls and -0.004 (s.e.=0.047) for boys. Also notable is the large difference between the estimates from the falsification regressions and from those obtained when the true treatment variable is used. The lack of any discerned effects when the placebo treatments are used provides further evidence suggesting that the estimated treatment effects are not capturing a spurious correlation between the proportion female and time-varying school factors. These results also suggest that the peer effects operate mainly at the grade level with no spillover effects on adjacent grades.

C. Heterogeneous Treatment Effects

To gain further insights on the extent of gender peer effects, we explore heterogeneous effects of the proportion of girls across different dimensions. In Table 4, we report heterogeneous treatment effects of the proportion female for two sub-samples stratified by the average years of schooling of both parents (average above or below 12 years of schooling) and

[26] Note that the number of observations is slightly different in columns 7-10 from the respective sample sizes in columns 2-3 and 5-6 because for a small number of schools in our sample there are no classes in one of these adjacent cohorts. We re-estimate the models reported in columns 2 and 5 using the same sample of columns 7-10. The results are virtually identical and are available upon request.

for a sub-sample of new immigrants (5 or less years since immigration).[27] The results clearly show that the positive impact of the proportion of girls in class is larger among students with lower parental education. The benefits are even higher (both in absolute terms and relative to the outcome means) for new immigrants. These results hold for both boys and girls. The larger effect among students from low socio-economic background implies a more dramatic impact of the proportion female than the one discussed above for the full sample and suggests that the benefits from having a higher proportion of female peers are larger for students who are likely to attend classes with higher levels of disruptions and violence and for students who are likely to need additional instruction time. We will return to the interpretation of these findings in the last section.

[Table 4]

Since the larger variability in the proportion of female students arises from small schools, we also examine whether the effects obtained for students' achievement arise from small schools only. Web Appendix Table 2 presents the results for samples stratified by school size: average cohort size below 200 (136 schools) and average cohort size of 200 or above (128 schools). Interestingly, the effects of the proportion of female students are very similar in the samples of small and large schools and are virtually identical to those obtained in the full sample. This result suggests that the benefits of a larger proportion of female students are evident for small as well as for large deviations around the school mean.

We also estimate models that allow for nonlinear effects of the proportion female on student outcomes. The effects appear to be nonlinear, with impacts that grow with the increase

---

[27] Students with missing values in parental education (4 percent of the total sample) are excluded from this analysis. The results are not sensitive to the inclusion of these students in the low or high education group. We also estimate heterogeneous treatment effects by stratifying the sample by father's or mother's schooling and we obtain very similar results to those based on the stratification by the mean of parental schooling and reported here.

in the proportion female. Similar to Hoxby (2000) the highest impacts are evident when girls are a majority in the class, in our case, when the proportion female exceeds 58 percent (see Web Appendix Tables 3 and 4, and Lavy and Schlosser (2007) for a discussion of these results).

D. Effect on Elementary and Middle School Outcomes

The test score data we have for elementary and middle schools pool together only two cohorts of 5^{th} and 8^{th} grades, respectively. Therefore, the identification of within school variations is less powerful in these samples. Nevertheless, albeit to a lesser extent, we do find positive effects of the proportion of girls on test scores. Table 5 presents the results for 5^{th} grade (columns 1-4) and 8^{th} grade (columns 5-8) based on the same two specifications we use to estimate the high school outcomes equations.[28] The estimates for math and science scores in 5^{th} grade show positive effects although standard errors are sometimes too large to obtain significant estimates. Estimates for Hebrew and English are small for both genders and are not statistically significant.

[Table 5]

To reduce measurement error and improve precision, we also estimate the effects on achievement using the average test scores in math and science and the average test scores in Hebrew and English. The estimated coefficients for the average of students' math and science scores in 5^{th} grade are significant for both genders. The effect size suggests that a 10 percentage-point increase in the proportion of female students increases average test scores of girls and boys in these particular subjects by 3.5 and 3.1 percent of a standard deviation, respectively. These effects are slightly larger than Hoxby (2000) who found that a 10 percentage-point

---

[28] Here, however, we cannot include school specific time trends as we only have two years of data for each school.

increase in the proportion female increases students' math tests scores by 1-2 percent of a standard deviation in Texas elementary schools. Estimates of the effects on the average of Hebrew and English scores are positive but not significant. It is noteworthy that while girls perform remarkably better than boys in Hebrew and English, the effect of the proportion of girls on students' performance is only visible in math and science, subjects where girls have little to no advantage compared to boys. This suggests that girls' peer effects do not operate solely through spillovers of peers' higher achievement—an issue we explore in detail in the next section.

Results for 8th grade (Table 5, columns 5-8) show a strong effect of the proportion of girls on girls' math and English test scores, with smaller positive effects for boys but with large standard errors. [29]

V. Identifying the Mechanisms of Gender Peer Effects

The results reported above show that both boys and girls exhibit higher achievements when they have more female peers in their class. In this section, we explore the potential mechanisms through which girls may affect their peers' academic achievement. One obvious mechanism could be the spillover of girls' achievement. However, it seems unlikely that all gains in achievement are generated solely by this channel since we find positive effects of the proportion of girls even in subjects where girls have lower achievement than boys (e.g., the number of credit units in scientific subjects in high school or math and science test scores in

[29] Our attempt to estimate heterogeneous treatment effects by parental education or immigration status in the primary and middle school sample did not produce a clear pattern, probably due to the lack of statistical power in these samples.

elementary schools).[30] This is also consistent with Hoxby and Weingarth (2005) who find that even after controlling for peers' lagged achievement, race, ethnicity, and income, a higher proportion of girls in the class leads to higher test scores for both genders.

In this section, we examine other possible channels using a rich set of behavioral outcomes based on responses to school questionnaires of middle and elementary school students and teachers. We first show in Web Appendix Table 5 that all indicators of the quality of the classroom environment, as described by the students, are highly correlated with students' academic performance even after controlling for school fixed effects and students background characteristics. For example, we find that lower levels of classroom disruption and violence, better inter-student relationships and a higher quality of interaction between teachers and students are all positively associated with students' tests scores. While we cannot provide a casual interpretation to these correlations, the results reported in this table suggest that students' assessments of their classroom environment have a high informational content and that these mechanisms, as pointed out in the educational literature, might play an important role in student's learning process.[31]

We are aware, of course, that we are not able to measure all the relevant mechanisms and we cannot rule out the possibility that other mechanisms are in place but the analysis presented in this section provides important insights regarding the possible mediating factors that drive the positive effect of girls on students' achievement.[32] Our hypothesis is that if the effects of girls

---

[30] It could still be the case that gender peer effects are enacted solely through girls' higher achievement if there are spillovers across subjects.

[31] See Jerome Freiberg (1999) and Barry Fraser (1998) for recent reviews of the educational research literature about the validity of students and teachers assessment of the classroom environment and their associations with students' achievement.

[32] A further limitation is that we cannot identify the causal effect of the mechanisms on outcomes because the former are numerous and we have only one potential instrument

are partially being driven by a particular mediating factor, observing a significant effect of the proportion female on this factor provides some evidence for the validity of this hypothesis.

In addition, regardless of their possible role as mediating factors, we believe that the effects of the proportion female on these behavioral outcomes are interesting in their own right, as exemplified by numerous studies that highlight their central role in school choice decisions (see, e.g., Caroline Hoxby 1998; Sandra E. Black 1999; Thomas J. Kane, Stephanie K. Riegg, and David O. Staiger 2006, and Berry J. Cullen, Brian A. Jacob and Steven Levitt, 2006) and in teachers transferring and quitting decisions (see e.g., Donald Boyd Hamilton Lankford, Susana Loeb, and James Wyckoff 2003; Hanushek, Kain, and Rivkin, 2004b).

A. ***Classroom Environment***

We focus on eight items in the student questionnaire that relate to the classroom and school environment. To obtain a more general picture of the possible mechanisms and to gain statistical power, we also group the eight outcomes into the following three categories: classroom disruption and violence; inter-student relationships; and teacher-student relationships. Low scores achieved in the first category and high scores achieved in the second and third categories point to improved outcomes.

Following Kling et al. (2007) we compute the average effect τ_c for each category c by averaging across the standardized effects of the individual items estimating a system of seemingly unrelated regressions.[33] As there is no prior information to justify a particular

---

[33] That is, we define the average effect of the proportion of female students for category c as $\tau_c = \frac{1}{k_c}\sum_{k=1}^{K_c}\frac{\pi_{kc}}{\sigma_{kc}}$ where k_c is the number of outcomes included in category c, π_{kc} is the effect on outcome k included in category c, and σ_{kc} is the standard deviation of the outcome. We treat (σ_{kc}) as known based on the results of Kling and Liebman (2004) and given that we have a large sample.

weighting, we assign equal weight to all outcomes within a category as this provides a more transparent interpretation.[34]

Table 6 reports within school estimates using pooled data of 5th and 6th graders, a second sample of 7th through 9th graders, and a pooled sample of 5th through 9th graders. We report results for individual outcomes as well as the average effect for each category. Similar to the estimations for the high school outcomes, we also estimate falsification regressions for all items in the student questionnaire. The results of these falsification tests are reported in Web Appendix Table 6, and show that all estimates of the placebo treatments are small, have inconsistent signs, and are not significantly different from zero.

[Table 6]

B. Classroom Disruption and Violence

The analysis on classroom disruption and violence is based on the following items: (1) "Frequently the classroom is noisy and not conducive to learning"; (2) "There are many fights among students in my classroom"; (3) "Sometimes I'm scared to go to school because there are violent students". The mean responses of girls and boys to these questions are almost identical as seen from columns 1-2 (elementary school) and columns 5-6 (middle school), implying that students' subjective assessment of the classroom environment is similar across both genders.

The estimates reported in columns 3-4 and 7-8 in the first panel of Table 6 suggest that a higher proportion of girls in a class significantly lowers the level of disruption and violence. This effect is evident in each of the three items, as reported by both boys and girls, and it is

[34]As an alternative strategy, we also construct aggregate outcomes by averaging across the standardized outcomes included in each category and estimate the effects of the proportion of female students on these aggregate outcomes. In practice, both methods provide identical estimates when there are no missing values in item responses and the model has no additional covariates besides the treatment variable. The results based on the simple averaged outcomes are virtually identical to the result that we report in Table 6.

equally precise and important in both elementary and middle school. In columns 9-10 we report the estimates from a sample that pools all grades together. The pooled sample provides some gain in precision, reducing the standard errors by 20-40 percent. For example, the estimate for the effect of the proportion of girls on students' reports regarding the level of noise in the classroom is -0.254 (s.e.=0.089) for girls and -0.218 (s.e.=0.080) for boys.

The average effect across the three items included in the classroom disruption and violence category is more precise than the estimates for the individual items: the estimate for girls in the pooled sample is -0.302 (s.e.=0.058) and for boys it is -0.233 (s.e.=0.049). Overall, these results suggest that having more girls in a class highly improves the learning and safety climate by lowering the disruptions during lessons, lowering the incidence of fights, increasing the safety of students, and lowering their anxiety about attending school. Beyond the direct effect, personal safety in school can also indirectly affect students' achievements by improving motivation, concentration, and other non-cognitive factors that are important for learning. In addition, fewer disruptions during class are likely to lead to a more efficient use of the instruction time.[35] Overall, the evidence about this mechanism is consistent with the findings reported in David Figlio (2007) and Joshua Kinsler (2007) that disruptive behavior of students has negative ramifications for their peers' test scores.

C. Inter-Student Relationships

Two items in the questionnaire ("I feel well adjusted socially in my class" and "Students in my class help each other") supply an indication of the quality of inter-student relationships

---

[35] We stratify the sample by low and high parental education and, consistent with our results on achievement, we find a stronger impact for students with low parental education. For the other mechanisms, there is not such a clear pattern, perhaps because the group of low parental education has higher mean rates of disruption and violence but report similar inter-student and teacher-student relationships. The estimates for new immigrants are much noisier and not significantly different from our main results.

that can be conducive or harmful to learning and achievement. Being well adjusted and accepted socially among classroom peers may improve a student's self-confidence, self-image, motivation, and other non-cognitive attributes that might be essential for effective learning. Cooperation between students may comprise help with homework or with test preparation, both of these implying additional instruction times and better learning.

Boys and girls have similar feelings regarding their social adjustment in class. On the other hand, girls have a more favorable view of the cooperation between students in a class, especially at the middle school level, suggesting perhaps that girls are more cooperative than boys are. The within school estimates show that a higher proportion of girls in a class improves both outcomes significantly. The effect among girls in elementary school is larger than among boys, but in middle school, it is equal for both genders. The estimated effects are larger in middle school, reflecting perhaps the increased importance of social interaction among teenagers and a more pronounced effect of girls in a more 'turbulent' classroom. The average treatment effect of these two items over all grades is 0.302 (s.e.=0.057) for girls and 0.155 (s.e.=0.049) for boys.

D. The Quality of Teacher-Student Relations

We used three items to examine the effect of the proportion of girls in a class on the relationships between students and teachers. The first item identifies how rude students are to their teachers ("Students frequently talk back to teachers"). The effects of the proportion of girls in a class are significant and negative, meaning that a higher proportion of girls lead to a lower frequency of offensive treatment of students towards teachers, with the effects being similarly reported by boys and girls. In contrast, we do find a different effect for boys and girls when we look at two other aspects that affect the quality of the relationships between students and teachers. For these two items "There are good relationships between teachers and students" and "There is mutual respect between teachers and students", the estimates are much higher for boys than for girls in both elementary and middle schools. Overall, we can conclude that the peer effect of girls in school is working through the quality of teacher-student relationships as well, and that it is doing so largely among boys.

E. Further Evidence on the Mechanisms from Students Fixed Effects Estimates

The structure of the GEMS allows us to follow a sample of students from elementary schools (at 5th or 6th grade in 2002, 2003) to middle schools (at grade 7th, 8th or 9th in 2004 or 2005).[36] We take advantage of this feature and construct a longitudinal dataset at the student level to examine how changes in students' assessments of their classroom environment are associated with changes in the proportion of female peers (due to their transition from elementary school to middle school).[37] We estimate the following first difference equation by differencing out two relationships like equation (1) for each student (one for middle school and one for elementary school):

$$(3)\ y_{igst}^{MS} - y_{igst}^{ES} = \alpha_g^{ES} + \alpha_g^{MS} + \beta_s^{ES-MS} + \gamma_t + x_{igst}^{'}\lambda_1 + S_{gst}^{'ES}\lambda_2 + S_{gst}^{'MS}\lambda_3 + \pi(P_{gst}^{MS} - P_{gst}^{ES}) + \Delta\varepsilon_{igst}$$

Where *ES* denotes elementary school and *MS* denotes middle school. A student fixed effect is differenced out from this equation and we add controls for students' background characteristics, the average characteristics of their cohort in elementary and middle school, and a

---

[36] Specifically, we follow students who were in 5th or 6th grade in 2002 or 2003 to their respective grades in 2004 and 2005. We are therefore able to follow about 50 percent of the sample of elementary school students reported in columns 1-4 of Table 6 finding roughly 80 percent of them. We do not link between datasets from consecutive years because almost all localities were sampled once every two years.

[37] We cannot perform a similar analysis with test scores as an outcome because the longitudinal panel data does not include students that are observed at both 5th and 8th grade, and these are the only grades at which the cognitive tests are administered.

grade fixed effect in elementary and in middle school. In addition, we add a fixed effect for all students who attended the same elementary and the same middle school ($\beta_s^{ES\_MS}$).[38] Identification is therefore achieved from variation in students' assessments of their classroom environment across elementary and middle school within students from adjacent cohorts who followed the same transition path.

Table 7 presents student fixed effects estimates of the effect of the proportion of girls on the learning and classroom environment. These estimates are remarkably similar to the school fixed effects estimates presented in columns 9-10 of Table 6, although they are less precise due to the lower variation. For example, focusing on the average effects, the estimates for the effect on classroom disruption and violence in the student's fixed effects model are -0.321 for girls and -0.195 for boys, while the respective school fixed effects estimates reported Table 6 are -0.302 and -0.233. We view these results as reassuring the informational content of the survey data. Furthermore, the fact that we obtain similar results when applying two distinct identification strategies provides additional supporting evidence for the causal interpretation of the effects of the proportion female on the classroom environment.[39]

[Table 7]

F. Teachers' Fatigue and Work Satisfaction

Complementary to the analysis on classroom environment is that of the impact of the proportion of female students on teachers' fatigue, burn-out, and work satisfaction. These

---

[38] The results are virtually identical when these controls are omitted from the regression. They are also qualitatively unchanged when we simply include a separate fixed effect for each primary and each middle school.

[39] As a further check on the informational content of the students questionnaires, we compare the correlation of students' assessment of their classroom environment when they are observed in the same school in two different grades (7th and 9th grade) with the correlation of students' assessments when they are observed in two different schools (due to the transition from elementary school to middle school). The results show a significantly higher correlation for students' assessments of the classroom environment when they are observed in two different grades in the same school.

factors are likely to affect teachers' motivation and productivity. To analyze these aspects we look at the GEMS teacher questionnaire that included the following three relevant items: (1) "I feel burned-out as a teacher"; (2) "I feel that I have too much workload"; (3) "I am satisfied with my work at school".

We were able to match the home classroom teachers to their students for the elementary and middle school data. However, the contact time between the home classroom teacher and her students in middle school is very limited, only a few hours a week, while in elementary school most of the classes are taught by the home classroom teacher, especially in the lower grades. We therefore focus in this analysis only on the sample of 17,529 home classroom teachers in 1^{st} to 6^{th} grades in 1,038 schools. Table 8 presents estimates of the effect of the proportion of girls on teachers' responses to the above three items. We report estimates from school fixed effects models that control for the mean characteristics of the grade and include grade and year dummies. These estimates are reported in panel A of the table.

[Table 8]

The mean of teachers' responses to the statement "I feel burned-out" is 2.6 on a scale of one (completely disagree) to six (strongly agree). About a quarter of the teachers agreed to some extent with this statement, suggesting that a non-negligible number of teachers feel exhausted from their jobs. The estimates in column 3 show that this emotional-physical status of teachers is strongly and negatively related to the proportion of girls in their classroom. The estimate based on the full sample is -0.265, and it is only marginally significant (t-value=-1.4). However, when the sample is limited to lower grades (1^{st} through 4^{th}), where the teachers are most likely to be teaching only the grade for which treatment is measured, the estimate increases significantly (-0.637 with a t-value=-2.6). A larger and more precise effect is also estimated for a sample that includes only math and language home teachers, who are also more likely to be teaching most of their hours in the grade where the treatment variable is measured.

In the lower panel of column 3, we report estimates from within school regressions where various measures of the classroom environment (as reported by students) replace (one at a time) the treatment variable of the proportion of girls in the grade. Not surprisingly, these estimates indicate that the "fatigue" of teachers is highly negatively correlated with the quality of the classroom environment: teachers feel much more exhausted when classrooms are noisy, when there are more fights among students, when students are more abrasive towards their teachers, and when students and teachers do not have good relationships and do not respect each other.

We cannot interpret these estimates as causal because there might be a third factor affecting both the classroom environment and teachers' fatigue or there may be reverse causality. However, these within school associations are consistent with the effects of gender composition on the classroom environment and therefore can be viewed as channels through which gender composition may affect teachers.[40] If teachers who feel burnt-out have lower productivity, it is reasonable to think that the positive effects of the proportion of female students on peer achievement is also driven by a lower level of teacher fatigue and burnout.

Columns 4-5 report the effects of gender classroom composition (panel A) and correlations with classroom environment (panel B) for the two remaining questionnaire items concerning teachers' workload and work satisfaction. Overall, neither outcome is affected by the proportion of girls in the grade, as the six estimates reported in panel A suggest. The associations between teachers' reports on having a high workload and the classroom

---

[40] In addition, these within school associations between the classroom environment and teachers' sense of fatigue provide further evidence for the high informational content contained in both teachers' and students' responses to the school survey.

environment shown in panel B have the expected sign but are only marginally significant and much smaller (only about a quarter) in magnitude compared to the associations reported for teachers' burnout. This may suggest that girls do not have much of an effect on teachers' workload since the latter is only weakly related to the classroom environment.

Despite teachers' complaints about feeling burnout and having a high workload, only 3 percent of the teachers reported low scores (1-3) for work satisfaction. This outcome has a very high average (5.456) and a much lower spread around the mean (s.d.=0.82), so that any effect will be harder to detect. Nevertheless, we do find that teachers' satisfaction with school is correlated with the classroom environment indicators, although these estimates are much smaller (about half the size) than the corresponding ones for teacher burnout. On the other hand, the proportion of female students has no effect on teachers' work satisfaction. The fact that there is no correlation between the teachers' workload and satisfaction indicators and the proportion of girls may be a result of other factors being more dominant than the proportion of girls in the determination of these two indicators such as the level of compensation and other duties at school.

VI. Change in the Classroom Gender Composition versus Change in Behavior

The results discussed above clearly show that a higher proportion of girls in a class leads to an improved learning environment, as reflected by a lower level of violence and classroom disruption and better inter-student and teacher-student relationships. But one central question remains: how much of the peer effect on the learning environment is due to changes in gender classroom composition and how much to changes in the behavior of students? Based on additional items in the student questionnaire we are able to provide a limited answer to this question. By contrasting the information students provided on how they view their classroom environment with their answers to questions about their own behavior, we find very sharp and informative differences.

Table 9 presents estimates of the effect of the proportion of girls in the classroom on items that measure (based on self-reporting) the student's understanding of the learning and discipline requirements in school, his/her involvement in fights with other students, and his/her relationship with the teachers. Similar to the previous section, we summarize the effects on the various outcomes related to student behavior by grouping them into broader categories and computing the average effect for each category.

[Table 9]

The striking overall pattern seen in this table is of no systematic or significant effect of the proportion of girls in the classroom on any of these measures of students' behavior. The most obvious example is the item on being involved in many fights at school during the current year. Boys are much more likely to be involved in fights than girls, with a mean score of 2.372 versus 1.490 on a scale of one (completely disagree) to 6 (strongly agree) in elementary school. However, the effect of the proportion of girls in a class in elementary school is positive and significant both for boys and girls and in middle school it is negative (though not significant) for boys and not significant for girls. Therefore, if there is any effect on violent behavior of students, it has an opposite sign to the effect on the classroom average. This suggests that the effect of the proportion of girls on disruption and violence is mainly driven by a change in the composition of the class and not by changes in individual behavior of students: as the number of girls in the class increases, so does the proportion of well-behaved students, and therefore the mean level of violence is reduced. On the other hand, there are no behavioral changes among girls or boys. This pattern of results is also seen from the estimates that are based on the longitudinal panel sample described in section VE. We present these student fixed effect estimates in Web Appendix Table 7.

Another potential behavioral change is that of study effort. The lower panel in Table 9 reports the effect of the proportion of girls on weekly homework hours in math, Hebrew, English, and science and technology. There is no systematic pattern in the 16 estimates (four subjects for each gender, in elementary and in middle school) in terms of sign or precision: some are negative and others positive and most are not significantly different from zero. We do observe that girls spend more time doing homework than boys in all subjects (0.73 hours more in elementary school and 0.83 hours more in middle school), but having more girls in a class has no effect on these outcomes, suggesting that the positive gender effects on scholastic achievement reported in section IV operate through channels other than an increase in learning effort. The fact that we do not find an effect on an out-of-school outcome (time at home spent on homework) may be viewed as another indication that it is not a cohort effect that drives the results reported in this and in the previous section.

We also estimate falsification or placebo regressions for all items examined in this section. The results are reported in Web Appendix Table 8. All these estimates are small, have inconsistent signs, and are not significantly different from zero.

VII. Conclusions

In this paper, we measure empirically the extent of peer effects of female students in elementary, middle, and high school on students' academic achievement. Using unique and rich data on behavioral outcomes, we are able to look into the "black box" and explore alternative mechanisms through which the classroom gender composition may affect students' performance. The data also allow us to disentangle two different channels through which these mechanisms may affect students, one that operates through a change in the gender composition of the classroom and a second that reflects changes in the individual behavior of students.

The evidence provided in this paper suggests that a higher proportion of female peers improves the scholastic achievements of both boys and girls. The effects seem to be larger at higher proportions of girls in the classroom, in particular, when girls are a majority in the cohort. The benefits from a higher proportion of female peers are higher among disadvantaged students or students from low socio-economic background as suggested by the larger impacts we found among new immigrants and students with low parental education. This result is consistent with the class size empirical literature that found a stronger impact of class size reductions among students from disadvantaged backgrounds (see e.g., Angrist and Lavy 1999 and Alan Krueger 1999) and with the theoretical prediction of Lazear (2001). Both, class size reductions and higher proportions of female students, can be seen as alternative treatments that reduce the amount of classroom disruptions and violence.

The positive impacts of the proportion female on academic outcomes do not appear to be generated entirely by spillover effects of girls' achievements. An exploration of the mechanisms of the gender peer effects shows that a higher proportion of females in a class leads to a better classroom and learning environment. Students who have more female peers report a lower level of classroom violence and disruption and better relationships with other students and with teachers. The effects on improved classroom environment appear to come from a change in the classroom composition and not from changes in students' individual behavior or in their study effort. The benefits from a higher proportion of girls in the classroom are also due to lower fatigue and burnout among teachers, which probably affects their productivity. We should note that part of the effect of girls on classroom climate could be due to the fact that most of the teachers are women. If male teachers were able to handle the boys better, having more male teachers could have lowered perhaps the role of girls in shaping classroom climate. Nevertheless, this observation does not make our results less relevant since there is a majority of female teachers in most countries.

The findings that both boys and girls excel in an environment with more girls and that there are large similarities across gender in the importance of the various mechanisms through which gender peer effects operate, complicate the social choices regarding single sex classes and schools. The gain for females from school or classroom gender segregation is offset by the loss for males. For example, placing girls in single-sex classes in math and sciences would deny boys the positive externalities of girls. Still, a complete analysis on the benefits or losses of single-sex versus coed schooling would require additional assumptions regarding the welfare function and explicit modeling of peer interactions and the educational production function.

Our results also provide direct evidence of the possible negative consequences of imbalanced sex ratios in co-educational schools that could emerge from a disproportionate increase in the number of single-sex classes for girls. A general implication of our findings is that the gender mix of a class should be taken into account in inter- and intra-school resource allocation decisions, especially when the proportion of girls is particularly low or in schools serving low SES populations. For example, school cohorts with a high proportion of boys could be allocated into smaller classes to offset the negative impact on the classroom environment. In addition, teachers who are better trained to deal with behavioral problems could be assigned to classes that have a higher proportion of boys. A further implication of our results is that the classroom gender composition could be taken into account in decisions regarding the placement of low achievers, new immigrants or students with behavioral problems.

While our research design does not permit us to identify the relative weight of each mechanism and does not rule out the possibility that other mechanisms might be at play, our results provide important insights towards the understanding of channels through which peers influence student learning that go beyond the focus on the gender dimension. For example, our findings highlight the importance of developing a richer battery of survey instruments to distinguish between peer effects that result from changes in individual behavior and peer effects that result from externalities on the classroom environment. Further evidence on the relative role of these two distinct channels would have important implications for equity and efficiency considerations in the placement of students across classes.

References

Ammermueller, Andreas, and Jorn-Steffen Pischke. 2009. "Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study." *Journal of Labor Economics*, 27(3): 315-348.

Angrist, Joshua, and Kevin Lang. 2004. "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program." *American Economic Review*, 94(5): 1613-1634.

Angrist, Joshua, and Victor Lavy. 1999. "Using Maimonides' Rule to Estimate the Effect of Class Size on Children's Academic Achievement." *Quarterly Journal of Economics*, 114(2): 533-575.

Angrist, Joshua and Victor Lavy. 2009. "The Effects of High Stakes High School Achievement Awards: Evidence from a School-Centered Randomized Trial." *American Economic Review*, 99(4): 1384-1414.

Arcidiacono, Peter and Nicholson Sean. 2005. "Peer Effects in Medical School." *Journal of Public Economics*, 89(2-3): 327-350.

Black, E, Sandra. 1999. "Do Better Schools Matter? Parental Valuation of Elementary Education." *Quarterly Journal of Economics*, 114(2): 577–99.

Boozer, Michael, and Cacciola Stephen. 2001. "Inside the 'Black Box' of Project Star: Estimation of Peer Effects Using Experimental Data." Yale University Economic Growth Center Discussion Paper 832.

Boyd, Donald, Hamilton Lankford, Susana Loeb, and James Wyckoff. 2003. "Analyzing the Determinants of the Matching of Public School Teachers to Jobs: Estimating Compensating Differentials in Imperfect labor markets." National Bureau of Economic Research Working Paper 9878.

Campbell, Patricia, and Jo Sanders. 2002. “Challenging the System: Assumptions and Data behind the Push for Single-Sex Schooling.” In *Gender In Policy and Practice: Perspectives on Single-Sex and Co-educational Schooling*, ed. Amanda Datnow and Lea Hubbard, 31-47. New York: Routledge Falmer.

Cullen, Berry J., Brian A. Jacob, and Steven Levitt. 2006. “The Effect of School Choice on Participants: Evidence from Randomized Lotteries.” *Econometrica*, 74(5): 1191-1230.

Figlio, David. 2007. “Boys named Sue: Disruptive Children and their Peers.” *Education Finance and Policy*, 2(4): 376-94.

Freiberg, H. Jerome. 1999. “School climate: Measuring, improving and sustaining healthy learning environments.” Philadelphia: Falmer Press.

Fraser, J. Barry. 1998. “Classroom Environment Instruments: Development, Validity, and Applications”, *Learning Environment Research*, 1(1): 7-34.

Gould, Eric, Victor Lavy, and Daniele Paserman.2009. “Long Term Classroom Peer Effects: Evidence from Random Variation in Enrollment of Disadvantaged Immigrants.” *Economic Journal*, 119(540):1243-1269.

Hanushek, Eric, John Kain, Jacob Markman, and Steven Rivkin. 2003. “Does Peer Ability Affect Student Achievement?.” *Journal of Applied Econometrics*, 18(5): 527-544.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. 2004. “Disruption versus Tiebout improvement: the Costs and Benefits of Switching Schools.” *Journal of Public Economics*, 88(9): 1721–1746.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. 2004. “Why Public Schools Lose Teachers.” *Journal of Human Resources* 39(2): 326-354.

Harwath, Irene, Mindi Maline, and Elizabeth Debra. 1997. "Women's colleges in the United States: History, Issues and Challenges." Washington, D.C: Office of Educational Research and Improvement at the U.S. Department of Education.

Hoxby, Caroline M. 1998. "When Parents Can Choose, What Do They Choose? The Effects of School Choice on Curriculum and Atmosphere." In *Earning And Learning: How Schools Matter*, ed. Susan. Mayer and Peter G. Peterson, 281-317. Washington, D.C: The Brookings Institution Press.

Hoxby, Caroline. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." National Bureau of Economic Research Working Paper Working Paper 7867.

Hoxby, Caroline M., and Gretchen Weingarth. 2005. *Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects*. Mimeo. Cambridge: Harvard University.

Kane, Thomas J., Stephanie K. Riegg, and Douglas O. Staiger. 2006. "School Quality, Neighborhoods and Housing Prices." *American Law and Economics Review*, 8(2): 183-212.

Kinsler, Joshua. 2007. *Suspending the Right to an Education or Preserving It? A Dynamic Equilibrium Model of Student Behavior, Achievement, and Suspension*. Mimeo. Rochester: University of Rochester.

Kling, Jeffrey, and Jeffrey B. Liebman. 2004. "Experimental Analysis of Neighborhood Effects on Youth." Princeton University, Department of Economics, Industrial Relations Section, Working Paper 483.

Kling, Jeffrey, Jeffrey B. Liebman, and Katz F. Lawrence. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica*, 75(1): 83-119.

Krueger, Alan. 1999. “Experimental Estimates of Education Production Functions.” *Quarterly Journal of Economics*, 114(2): 497-532.

Lavy, Victor, and Analía Schlosser. 2005. “Targeted Remedial Education for Underperforming Teenagers: Costs and Benefits.” *Journal of Labor Economics,* 23(4): 839-874.

Lavy, Victor, and Analía Schlosser. 2007. “Mechanisms and Impacts of Gender Peer Effects at School.” National Bureau of Economic Research Working Paper 13292.

Lazear, Edward. 2001. “Education Production.” *Quarterly Journal of Economics*, 116(3): 777-803.

Manski, Charles. 1993. “Identification of Endogenous Social Effects: The Reflection Problem.” *Review of Economic Studies*, 60(3): 531-542.

Morse, Susan. 1998. *Separated by Sex: A critical Look at Single-Sex Education for Girls*. Washington: The American Association of University Women Educational Foundation.

Rhodes, Virginia. 2005. “Kids on the move: The Effects of Student Mobility on NCLB School Accountability Ratings”, *Perspectives on Urban Education*, 3(3): 1-30.

Sacerdote, Bruce. 2001. “Peer Effects with Random Assignment: Results for Dartmouth Roommates.” *Quarterly Journal of Economics*, 116(2): 681-704.

U.S. Department of Education, Office of Planning, Evaluation and Policy Development Policy and Program Studies Service. 2005. *Single-Sex Versus Secondary Schooling: A Systematic Review*. Washington, D.C: U.S. Department of Education.

U.S. General Accounting Office. 1994. *Elementary School Children: Many Change Schools Frequently, Harming Their Education*. Washington, D.C: U.S. General Accounting Office.

Whitmore, Diane. 2005. "Resource and Peer Impacts on Girls' Academic Achievement: Evidence from a Randomized Experiment." *American Economic Review Papers and Proceedings*, 95(2): 199-203.

Zimmerman, David J. 2003. "Peer Effects in Academic Outcomes: Evidence from a Natural Experiment." *Review of Economics and Statistics*, 85(1): 9-23.

Table 1. Balancing Tests for the Proportion of Female Students

	Secular and religious elementary schools		Secular middle schools		Secular high schools		
Dependent variable	OLS (1)	School fixed effects (2)	OLS (3)	School fixed effects (4)	OLS (5)	School fixed effects (6)	School fixed effects + school time trends (7)
Father's years of schooling	-0.057	-0.245	0.858	0.184	0.606	0.517	-0.097
	(0.481)	(0.228)	(1.053)	(0.361)	(0.648)	(0.445)	(0.414)
Mother's years of schooling	-0.206	-0.283	0.305	-0.647	0.539	0.372	-0.133
	(0.476)	(0.236)	(0.992)	(0.435)	(0.597)	(0.412)	(0.403)
Number of siblings	-0.329	0.023	-0.239	-0.379	0.356	0.275	0.012
	(0.155)	(0.077)	(0.342)	(0.315)	(0.220)	(0.294)	(0.254)
New immigrant	0.015	0.006	-0.062	0.002	-0.115	-0.023	0.036
	(0.008)	(0.006)	(0.030)	(0.012)	(0.036)	(0.034)	(0.021)
Ethnicity							
Israel	-0.003	0.007	-0.063	-0.022	0.092	-0.042	-0.059
	(0.043)	(0.022)	(0.109)	(0.029)	(0.047)	(0.041)	(0.030)
Asia/Africa (excluding Ethiopia)	-0.047	-0.014	0.120	-0.007	0.027	0.028	0.038
	(0.023)	(0.015)	(0.055)	(0.021)	(0.037)	(0.025)	(0.025)
Ethiopia	0.013	0.012	-0.016	-0.002	-0.002	-0.005	-0.001
	(0.022)	(0.009)	(0.019)	(0.009)	(0.003)	(0.006)	(0.004)
Europe/America (excluding FSU)	-0.016	-0.026	0.054	0.032	0.082	-0.014	-0.002
	(0.022)	(0.012)	(0.050)	(0.020)	(0.027)	(0.015)	(0.018)
Former Soviet Union (FSU)	0.054	0.021	-0.095	-0.001	-0.199	0.033	0.024
	(0.041)	(0.014)	(0.141)	(0.027)	(0.069)	(0.055)	(0.028)
Enrollment	3.778	7.115	21.553	-3.908	24.449	-6.418	-34.673
	(7.828)	(3.039)	(40.124)	(14.093)	(47.509)	(28.052)	(20.691)

Notes: The table reports OLS and school fixed effects estimates from separate regressions of the relevant dependent variable on the proportion of female students. All regressions include year dummies. Regressions in columns 1-4 include also grade dummies. Regressions in even columns include also school fixed effects. Regressions in column 7 include school fixed effects and school specific linear time trends. Robust standard errors clustered at the school level are reported in parentheses.

Table 2. The effect of the Proportion Female on Student's School Mobility

	Secular and religious elementary schools (1st grade)		Secular middle schools (7th grade)		Secular high schools (10th grade)	
	Females	Males	Females	Males	Females	Males
	(1)	(2)	(3)	(4)	(5)	(6)
Outcome means						
Left the school in t+1	0.075	0.082	0.049	0.056	0.066	0.097
Regression estimates						
Left the school in t+1	-0.028	-0.002	0.024	0.028	-0.023	-0.001
	(0.029)	(0.032)	(0.041)	(0.045)	(0.024)	(0.030)

Notes: The table reports means of the dependent variable (first row) and estimates (second row) for the effects of the proportion of female students in a grade on the likelihood that a student left the school in the following year (by moving to another school, dropping out, etc.). Further details on the definition of the dependent variable are provided in the text. The data for columns 1 and 2 include 1st grade students in 2002. The data for columns 3 and 4 include 7th grade students in 2001 and 2002. The data for column 5 and 6 includes 10th grade students in 1993 through 2000. All regressions include controls for student covariates (mother's and father's years of schooling, number of siblings, indicators for immigration status and ethnic origin, and indicators for missing values in these covariates), cohort mean controls (students individual controls averaged by school and year), and a quadratic function of annual school enrollment. Regressions in columns 1 and 2 include also indicators for school type (secular/religious). Regressions in columns 3 and 4 also include school fixed effects and year effects. Regressions in columns 5 and 6 include also school fixed effects, year effects, and school linear time trends. Robust standard errors clustered at the school level are reported in parentheses.

Table 3. Estimates of the Effect of Proportion Female on Scholastic Outcomes in High School

	Main Results						Falsification Tests			
	Females			Males			Females		Males	
	Outcome means	Proportion female in the cohort		Outcome means	Proportion female in the cohort		Prop. female in t-1	Prop. female in t+1	Prop. female in t-1	Prop. female in t+1
Outcome	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Average Score	69.2	6.314	5.747	63.1	7.918	7.966	3.117	-1.686	0.220	-0.040
		(2.142)	(2.423)		(2.702)	(2.817)	(2.249)	(2.564)	(2.920)	(2.644)
Matriculation status	0.619	0.099	0.099	0.523	0.049	0.048	0.021	0.028	0.016	-0.004
		(0.041)	(0.044)		(0.045)	(0.047)	(0.043)	(0.046)	(0.045)	(0.047)
Number of credit units	20.6	1.455	1.415	19.2	1.389	1.456	0.369	0.080	-0.663	0.410
		(0.855)	(0.940)		(1.050)	(1.089)	(0.871)	(0.898)	(1.112)	(0.985)
Number of advanced level subjects in science	0.581	0.141	0.152	0.619	0.227	0.233	0.026	-0.001	-0.018	0.012
		(0.072)	(0.071)		(0.076)	(0.078)	(0.075)	(0.066)	(0.075)	(0.074)
Matriculation diploma that meets university requirements	0.559	0.086	0.089	0.473	0.084	0.081	0.006	0.018	0.036	0.026
		(0.047)	(0.048)		(0.046)	(0.048)	(0.039)	(0.043)	(0.042)	(0.045)
Year effects		✓	✓		✓	✓	✓	✓	✓	✓
School Fixed Effects		✓	✓		✓	✓	✓	✓	✓	✓
School Time Trend		✓	✓		✓	✓	✓	✓	✓	✓
Enrollment (2nd Poly.)		✓			✓		✓	✓	✓	✓
Individual Pupil Controls		✓			✓		✓	✓	✓	✓
Cohort Mean Controls		✓			✓		✓	✓	✓	✓
Number of students	205,891	205,891	205,891	199,038	199,038	199,038	201,374	205,351	194,691	198,224
Number of schools	264	264	264	264	264	264	254	262	254	262

Notes: The table reports means of the dependent variables (columns 1 and 4), estimates for the effects of the proportion of female students in a grade on students achievement in high school (columns 2, 3, 5, and 6) and falsification tests using the proportion of female students of cohort in t-1 (columns 7 and 9) or in t+1 (columns 8 and 10). Proportion female is measured in 10th grade. Individual controls include: both parents' years of schooling, number of siblings, indicators for immigration status and ethnic origin, and indicators for missing values in these covariates. Cohort mean controls include students individuals controls averaged by school and year. The regressions include school fixed effects and school specific linear time trends. Robust standard errors clustered at the school level are reported in parentheses.

Table 4. Heterogeneous Effects by Parental Education and Immigration Status

Outcome	Females: Full sample (1)	Females: Low parental education (2)	Females: High parental education (avg.>12) (3)	Females: New immigrants (4)	Males: Full sample (5)	Males: Low parental education (6)	Males: High parental education (avg.>12) (7)	Males: New immigrants (8)
Average score	6.314	8.168	3.113	34.066	7.918	11.509	3.168	5.897
	(2.142)	(2.639)	(2.960)	(11.278)	(2.702)	(3.628)	(2.822)	(13.697)
	69.18	*65.43*	*75.33*	*53.62*	*63.10*	*58.21*	*70.52*	*47.95*
Matriculation status	0.099	0.135	0.032	0.221	0.049	0.050	0.017	0.233
	(0.041)	(0.055)	(0.048)	(0.138)	(0.045)	(0.063)	(0.055)	(0.142)
	0.619	*0.518*	*0.769*	*0.481*	*0.523*	*0.407*	*0.679*	*0.401*
Number of credit units	1.455	2.142	0.022	8.496	1.389	1.328	0.420	3.967
	(0.855)	(1.114)	(1.083)	(3.834)	(1.050)	(1.416)	(1.123)	(4.815)
	20.61	*18.78*	*23.49*	*16.55*	*19.18*	*16.67*	*22.75*	*15.65*
Number of advanced level subjects in science	0.141	0.101	0.110	0.255	0.227	0.120	0.319	0.209
	(0.072)	(0.073)	(0.112)	(0.249)	(0.076)	(0.085)	(0.108)	(0.269)
	0.581	*0.369*	*0.883*	*0.469*	*0.619*	*0.375*	*0.941*	*0.507*
Matriculation diploma that meets university requirements	0.086	0.097	0.047	0.259	0.084	0.078	0.054	0.083
	(0.047)	(0.058)	(0.053)	(0.134)	(0.046)	(0.061)	(0.055)	(0.125)
	0.559	*0.437*	*0.736*	*0.385*	*0.473*	*0.343*	*0.648*	*0.315*
Number of students	205,891	115,949	82,577	13,729	199,038	107,616	83,665	12,787

Notes: The table reports heterogeneous effects by parental education and immigration status of the proportion of female students on matriculation outcomes. The table also reproduces the estimates from the full sample reported in columns 2 and 5 of table 3. The regressions control for students background characteristics and school time varying controls detailed in table 3. The regressions also include school and year fixed effects and school specific linear time trends and control for a quadratic function of enrollment. The high parental education sample includes all students with average education of both parents above 12 years of schooling. The total number of students in the two subgroups by parental education is lower than the number of students in the full sample since some students have missing values in parental education. The new immigrants sample includes all students who have lived in Israel for 5 years or less. Robust standard errors clustered at the school level are reported in parentheses. Outcome means are reported in Italics.

Table 5. Estimates of the Effect of Proportion Female on Scholastic Outcomes in Elementary and Middle Schools

	5th grade				8th grade			
	Females		Males		Females		Males	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Math	0.366	0.366	0.218	0.126	0.773	0.778	0.360	0.483
	(0.155)	(0.158)	(0.159)	(0.158)	(0.282)	(0.278)	(0.283)	(0.288)
Science and Technology	0.301	0.308	0.432	0.338	-0.088	0.071	-0.190	0.050
	(0.169)	(0.170)	(0.167)	(0.166)	(0.307)	(0.313)	(0.329)	(0.350)
Hebrew	0.078	0.094	0.131	0.031	0.335	0.287	0.031	0.051
	(0.148)	(0.150)	(0.157)	(0.158)	(0.249)	(0.261)	(0.326)	(0.333)
English	0.077	0.112	-0.088	-0.141	0.540	0.607	0.295	0.370
	(0.172)	(0.173)	(0.156)	(0.156)	(0.229)	(0.234)	(0.260)	(0.273)
Average score in Math and Science	0.350	0.343	0.310	0.212	0.327	0.428	0.123	0.307
	(0.135)	(0.138)	(0.142)	(0.142)	(0.256)	(0.256)	(0.279)	(0.288)
Average score in Hebrew and English	0.098	0.119	0.065	-0.024	0.390	0.401	0.174	0.213
	(0.132)	(0.134)	(0.132)	(0.132)	(0.200)	(0.208)	(0.255)	(0.264)
Year effects	✓	✓	✓	✓	✓	✓	✓	✓
School Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓
Enrollment (2nd Poly.)	✓		✓		✓		✓	
Individual Pupil Controls	✓		✓		✓		✓	
Cohort Mean Controls	✓		✓		✓		✓	
Number of students	56,288	56,288	57,527	57,527	52,551	52,551	53,042	53,042
Number of schools	999	999	999	999	389	389	389	389

Notes: The table reports school fixed effects estimates for the effects of the proportion of female students in a grade on students standardized tests scores in 5th (columns 1-4) and 8th (columns 5-8) grade. Individual controls include: both parents' years of schooling, number of siblings, immigration status, ethnic origin, and indicators for missing values in these covariates. Cohort mean controls include students individuals controls averaged by school and year. Robust standard errors clustered at the school level are reported in parentheses.

Table 6. Estimates of the Effect of Proportion Female on the Classroom Environment

	Secular and religious elementary schools (5^{th} and 6^{th} grades)				Secular middle schools (7^{th} through 9^{th} grades)				Full sample (5^{th} through 9^{th})	
	Outcome means		School fixed effects		Outcome means		School fixed effects		School fixed effects	
	Females (1)	Males (2)	Females (3)	Males (4)	Females (5)	Males (6)	Females (7)	Males (8)	Females (9)	Males (10)
Classroom disruption and violence										
1 Frequently the classroom is noisy and not conducive to learning	4.772	4.807	-0.318	-0.202	4.957	4.883	-0.211	-0.297	-0.254	-0.218
			(0.112)	(0.102)			(0.146)	(0.130)	(0.089)	(0.080)
2 There are many fights among students in my classroom	3.540	3.612	-0.707	-0.617	3.080	3.177	-0.594	-0.391	-0.669	-0.525
			(0.138)	(0.136)			(0.192)	(0.191)	(0.114)	(0.111)
3 Sometimes I'm scared to go to school because there are violent students	1.894	1.830	-0.328	-0.278	1.501	1.662	-0.175	-0.175	-0.247	-0.239
			(0.100)	(0.092)			(0.093)	(0.124)	(0.071)	(0.075)
Average effect			**-0.332**	**-0.253**			**-0.266**	**-0.212**	**-0.302**	**-0.233**
			(0.070)	**(0.060)**			**(0.094)**	**(0.081)**	**(0.058)**	**(0.049)**
Inter-student relationships										
4 I feel well adjusted socially in my class	5.181	5.196	0.234	-0.020	5.149	5.072	0.368	0.312	0.293	0.097
			(0.079)	(0.076)			(0.120)	(0.102)	(0.068)	(0.060)
5 Students in my class help each other	4.560	4.421	0.391	0.146	4.152	3.854	0.506	0.588	0.440	0.316
			(0.101)	(0.101)			(0.160)	(0.145)	(0.088)	(0.085)
Average effect			**0.260**	**0.048**			**0.360**	**0.336**	**0.302**	**0.155**
			(0.066)	**(0.061)**			**(0.103)**	**(0.081)**	**(0.057)**	**(0.049)**
Teacher-student relationships										
6 Students frequently talk back to teachers	3.969	4.026	-0.352	-0.370	4.490	4.364	-0.112	-0.173	-0.240	-0.282
			(0.143)	(0.135)			(0.166)	(0.163)	(0.109)	(0.105)
7 There are good relationships between teachers and students	4.523	4.392	0.098	0.262	3.792	3.640	0.235	0.410	0.153	0.326
			(0.104)	(0.112)			(0.164)	(0.158)	(0.090)	(0.091)
8 There is mutual respect between teachers and students	4.530	4.345	0.178	0.190	3.765	3.601	0.119	0.442	0.158	0.293
			(0.103)	(0.107)			(0.163)	(0.152)	(0.088)	(0.088)
Average effect (sign of item 6 is reversed)			**0.161**	**0.199**			**0.128**	**0.251**	**0.146**	**0.220**
			(0.080)	**(0.074)**			**(0.118)**	**(0.097)**	**(0.067)**	**(0.059)**
Number of students	105,590	107,803	105,590	107,803	135,826	135,031	135,826	135,031	241,416	242,834
Number of schools	1,010	1,010	1,010	1,010	395	395	395	395	1,302	1,302

Notes: The table reports means of the dependent variables (columns 1-2 and 5-6) and school fixed effects estimates for the proportion of female students on the classroom environment. The table also reports the average effect for the outcomes included in each category. The regressions control for student background characteristics (both parents' years of schooling, number of siblings, immigration status, ethnic origin and indicators for missing values in these covariates), cohort mean characteristics (students individuals controls averaged by school and year), a quadratic function of enrollment, year and grade dummies, and school fixed effects. Robust standard errors clustered at the school level are reported in parentheses.

Table 7. Student's Fixed Effects Estimates of the Effect of Proportion Female on the Classroom Environment

		Females (1)	Males (2)
Classroom disruption and violence			
1	Frequently the classroom is noisy and non conducive to learning	-0.337	-0.230
		(0.214)	(0.222)
2	There are many fights among students in my classroom	-0.838	-0.486
		(0.278)	(0.293)
3	Sometimes I'm scared to go to school because there are violent students	-0.323	-0.271
		(0.195)	(0.230)
	Average effect	**-0.321**	**-0.195**
		(0.107)	**(0.091)**
Inter-students relationships			
4	I feel well adjusted socially in my class	0.399	-0.045
		(0.189)	(0.191)
5	Students in my class help each other	0.455	0.058
		(0.210)	(0.256)
	Average effect	**0.301**	**0.001**
		(0.111)	**(0.116)**
Teachers-students relationships			
6	Students frequently talk back to teachers	-0.386	-0.388
		(0.255)	(0.245)
7	There are good relationships between teachers and students	0.202	0.326
		(0.204)	(0.253)
8	There is mutual respect between teachers and students	0.345	0.403
		(0.214)	(0.229)
	Average effect (sign of item 6 is reversed)	**0.211**	**0.229**
		(0.115)	**(0.112)**
Number of students		43,584	42,785
Number of elementary schools		982	980

Notes: The table reports estimates of the effect of the change in the proportion female (in the transition from elementary school to middle school) on the change in students assessment of their classroom environment. The sample includes 5th and 6th grade students in 2002 and 2003 and their follow up to middle school in 2004 and 2005. The boys sample is slightly smaller than the girls sample since boys who attend yeshiva middle schools are not included in the analysis. The regressions control for student's background characteristics, cohort mean characteristics and a quadratic function of enrollment in both elementary and middle school, year and grade dummies, and elementary x middle school fixed effects. Robust standard errors clustered by the interaction between elementary and middle school are reported in parentheses.

Table 8. Estimates of the Effect of Proportion Female on Teachers' Fatigue and Job Satisfaction

	Number of teachers (1)	Number of schools (2)	I feel burned-out (3)	I feel that I have too much workload (4)	I am satisfied with my work at school (5)
means	17,529	1,038	2.564	4.180	5.456
(s.d.)			(1.488)	(1.472)	(0.817)
A. Effects of the proportion of female students (grades 1st through 6th)					
Full sample	17,529	1,038	-0.265	-0.017	0.006
			(0.188)	(0.176)	(0.092)
Math & grammar teachers	16,837	1,037	-0.380	-0.039	0.032
			(0.193)	(0.178)	(0.094)
1st through 4th grade teachers	10,611	1,030	-0.637	-0.180	-0.002
			(0.244)	(0.238)	(0.117)
B. Within school associations with classroom environment (grades 5th and 6th)					
Frequently the classroom is noisy and not conducive to learning	6,844	1,001	0.238	0.054	-0.158
			(0.041)	(0.043)	(0.022)
There are many fights among students in my classroom	6,844	1,001	0.150	0.074	-0.091
			(0.030)	(0.031)	(0.017)
Students frequently talk back to teachers	6,844	1,001	0.190	0.056	-0.091
			(0.030)	(0.033)	(0.017)
There are good relationships between teachers and students	6,844	1,001	-0.332	-0.079	0.180
			(0.042)	(0.041)	(0.023)
There is mutual respect between teachers and students	6,844	1,001	-0.345	-0.087	0.179
			(0.043)	(0.042)	(0.024)

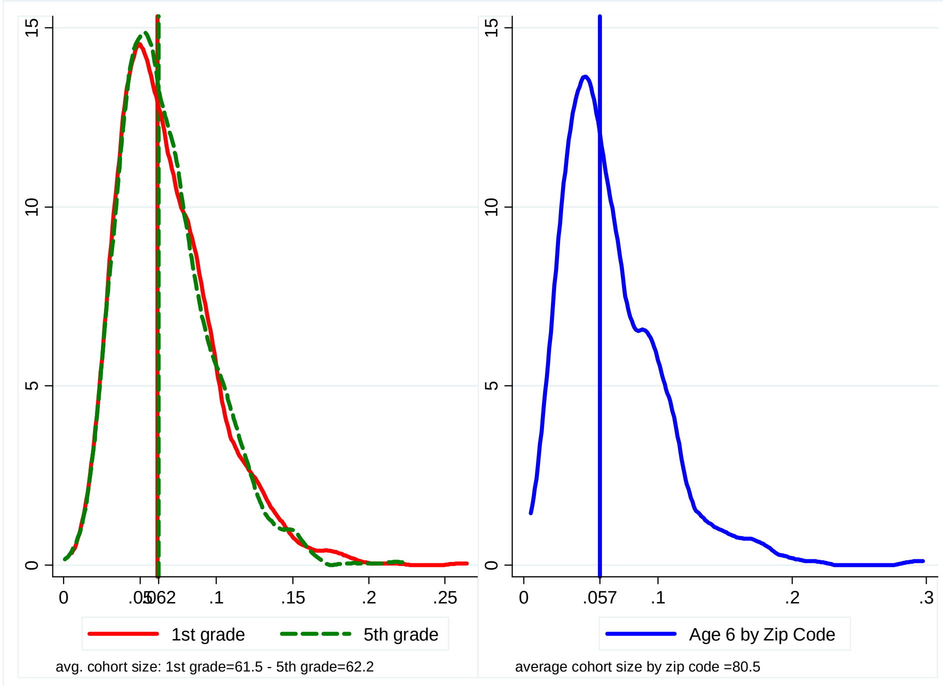
Notes: Rows 1 and 2 report means and standard deviations of teachers responses on different aspects concerning their work at school. Panel A reports the effects of the proportion of female students in a grade on teachers outcomes. Panel B reports within school associations between classroom environment (as reported by the students) and teachers' outcomes. Regression estimates are from models that include the control variables specified in Table 6. Robust standard errors clustered at the school level are reported in parentheses.

Table 9. Estimates of the Effect of Proportion Female on Student's Behavior

	Secular and religious elementary schools (5th and 6th grades)				Secular middle schools (7th through 9th grades)				Full sample (5th through 9th)	
	Outcome means		School fixed effects		Outcome means		School fixed effects		School fixed effects	
	Females	Males	Females	Males	Females	Males	Females	Males	Females	Males
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Self-discipline										
1 I understand well my teacher's scholastic requirements	5.027	5.016	0.038	-0.081	4.810	4.749	0.024	0.109	0.048	-0.005
			(0.067)	(0.068)			(0.105)	(0.106)	(0.058)	(0.058)
2 I know what behavior is allowed or forbidden in school	5.831	5.687	0.029	-0.047	5.638	5.426	0.024	0.070	0.035	-0.006
			(0.033)	(0.050)			(0.062)	(0.083)	(0.031)	(0.044)
3 This year I was involved in many fights	1.490	2.372	0.169	0.296	1.316	2.082	-0.093	0.076	0.060	0.228
			(0.071)	(0.102)			(0.079)	(0.136)	(0.053)	(0.082)
4 Sometimes the teachers treat me badly	2.680	2.946	0.150	0.143	2.989	3.189	0.251	-0.279	0.206	-0.019
			(0.131)	(0.122)			(0.171)	(0.169)	(0.104)	(0.101)
5 When I have a problem at school there is always someone I can turn to (from the teaching staff)	5.031	4.790	-0.030	0.074	4.591	4.234	-0.179	0.339	-0.080	0.173
			(0.107)	(0.111)			(0.171)	(0.179)	(0.093)	(0.096)
Average effect (signs of items 3,4 are reversed)			**-0.037**	**-0.074**			**-0.022**	**0.098**	**-0.029**	**-0.009**
			(0.051)	**(0.049)**			**(0.077)**	**(0.072)**	**(0.043)**	**(0.041)**
Study Efforts										
6 Weekly hours spent on homework in Math	3.337	3.144	-0.004	0.004	3.201	2.886	0.144	0.106	0.086	0.056
			(0.101)	(0.111)			(0.150)	(0.160)	(0.085)	(0.091)
7 Weekly hours spent on homework in Hebrew	2.546	2.371	-0.011	-0.010	1.970	1.812	0.140	0.099	0.049	0.006
			(0.110)	(0.108)			(0.166)	(0.153)	(0.092)	(0.087)
8 Weekly hours spent on homework in English	3.213	2.947	-0.046	0.025	2.917	2.621	0.157	0.266	0.045	0.109
			(0.110)	(0.117)			(0.153)	(0.166)	(0.089)	(0.095)
9 Weekly hours spent on homework in Science and Technology	2.445	2.395	0.137	-0.030	1.927	1.893	0.015	-0.132	0.102	-0.078
			(0.110)	(0.111)			(0.156)	(0.164)	(0.092)	(0.092)
Average effect			**0.014**	**-0.002**			**0.086**	**0.057**	**0.052**	**0.015**
			(0.059)	**(0.058)**			**(0.083)**	**(0.081)**	**(0.049)**	**(0.047)**
Number of students	105,590	107,803	105,590	107,803	135,826	135,031	135,826	135,031	241,416	242,834
Number of schools	1,010	1,010	1,010	1,010	395	395	395	395	1,302	1,302

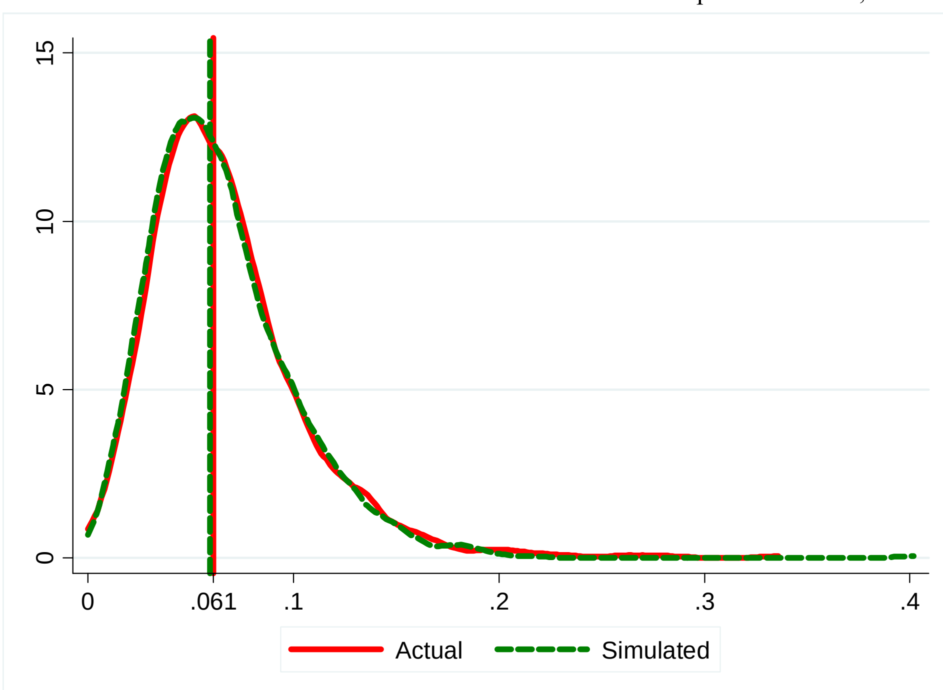
Notes: The table reports means of the dependent variables (columns 1-2 and 5-6) and school fixed effects estimates for the proportion of female students on students self-reported behavior and study efforts. Regression estimates are from models that include the control variables specified in Table 6. Robust standard errors clustered at the school level are reported in parentheses.

Figure 1a. Variation in the Proportion Female within Schools and Neighborhoods, 2002-2006



Notes: The left figure shows the distribution of the school standard deviation in the proportion female in 1st grade (solid line) and in 5th grade (dotted line). The sample includes the population of students in all public schools over the years 2002-2006. The right figure shows the distribution of the standard deviation in the proportion female aged 6 within geographical areas defined by the first 3 digits of the residential zip code. The sample includes all children aged 6 over the years 2002-2006 and was limited to areas where the size of a birth cohort lied within the range of the size of a school cohort. The variation in the proportion female within a geographical area is smaller than the within school variation because the avg. size of the cohorts is larger in the geographical areas. Vertical lines indicate the median of each distribution.

Figure 1b. Actual and Simulated within School Standard Deviation in the Proportion Female, Elementary Schools



Notes: The figure shows the standard deviation in the proportion female for each elementary school included in the analysis sample (solid line) and the simulated standard deviation in the proportion female for each of the schools (dotted line). Vertical lines indicate the median of each distribution. Details about the Monte Carlo simulation are provided in the text.