CESifo Working Papers

www.cesifo.org/wp

Conditional Gender Peer Effects?

Jon Marius Vaag Iversen
Hans Bonesrønning

CESifo Working Paper No. 4565
Category 5: Economics of Education
January 2014

CESifo *Working Paper No. 4565*

Conditional Gender Peer Effects?

Abstract

While the current empirical literature on peer group effects in schools highlights that credible causal peer effects cannot be estimated unless parental sorting is taken into account, the present paper highlights that causal peer effects might be conditional on the learning environment in which they occur. This approach is motivated by the existing theoretical literature which indicates that peer effects cannot be estimated without taking into account the role of school decision makers. We present indicative empirical evidence that gender peer effects in the Norwegian elementary school are conditional upon the level of special education provided.

JEL-Code: I200, I280.

Keywords: peer effects, conditional causal effects, special education.

Jon Marius Vaag Iversen
Trøndelag Research Institute
Norway - 7491 Trondheim
jon.iversen@svt.ntnu.no

Hans Bonesrønning
Norwegian University of Science and Technology
Dragvoll, 7/ 5/ 7564
Norway - 7491 Trondheim
hansbo@svt.ntnu.no

1. Introduction

Many education production function analyses, the 1966 Coleman report being the first, provide evidence that individual student achievement is affected by the peers in the classroom. Usually, peer effects are estimated while holding teacher characteristics, and the other factors that are assumed to affect student performance, constant. Most of these analyses do not take into account that peer effects might interact with other characteristics of the school organization- even though it seems quite likely that such interactions are important. For instance, causal peer effects might be conditional upon the average teacher quality, or in general, on the resources and institutions that characterize the educational system under scrutiny.

The purpose of the present paper is to contribute to the discussion about conditional, or context-dependent, peer effects. We do this by investigating whether the magnitude of *gender* peer effects depends on strategic actions from school authorities, notably we focus on the use of special education resources. Gender peer effects might depend on special education resources if, for instance, such resources reduce the congestion of Lazear-type congested classrooms (Lazear, 2001). The simple mechanism might be that less teaching time is lost when misbehaving male students are treated with special education resources. The hypothesis to be investigated is thus that the magnitude of gender peer effects decreases when special education increases. The Norwegian elementary school system makes up the context, and two of the established institutions herein might be of particular importance. First, misbehaving students have the legal right to special education resources, implying that disproportionate proportions of special education resources are allocated to boys. Second, compulsory schools are required to practice a full inclusion policy, implying that special education is offered within ordinary classrooms.

The present paper takes the insights from a small number of theoretical contributions about *school* decision making (Lazear (2001) being one of these) to the recent empirical peer group literature that highlights the endogeneity problems originating from parental sorting. Thus, we treat the student body composition as a result of parental sorting into schools, and the amount of special education resources as a result of the school decision makers' responses to the student body composition. To deal with the econometric challenges associated with parental sorting into schools we follow the existing empirical literature (see for instance Hoxby (2000), Lavy and Schlosser (2010), Black, Deveraux and Salvanes (2010) and Bifulco, Fletcher and Ross (2011)), and exploit idiosyncratic variations in gender composition across adjacent cohorts of students within schools.

Most of the analyses are devoted to the school decision makers' responses to the student body composition. Initially, we lay the ground for the rigid analyses by providing evidence that the proportion of eligible students in a grade-in-school correlates with the proportion of boys (our peer group measure) in the same grade-in-school, and that, after controlling for observable student characteristics, there is a significant and substantial increase in the proportion of eligible students from one year to the next in the period 2007-2009, probably reflecting that the introduction of national tests in 2004 -2007 "shocked" the system. Performing an education production function analysis with school fixed effects we find no significant gender peer effects for all students for the entire 2007-2009 time period. However, year-by-year estimations of the education production function shows that the gender peer effects are quite large and highly significant in 2007 and small and insignificant in 2009. The rest of the analyses put these basic findings under scrutiny.

The major challenge for the analyses is that the learning environment, to the extent that is determined by special education resource allocations, is endogenous and determined in complex interactions between actors in the education production process: parents, teachers, school principals, and psychological expertise located outside the schools, but inside the municipalities all have a say on these decisions. Our basic estimation strategy is to take advantage of the large increase in the amount of special education resources in the time period 2007-2009. We investigate whether these changes had consequences for the gender peer effects by using the variation in gender composition and special education across three adjacent cohorts, getting rid of much of the year-to-year variation in special education that is due to student and teacher idiosyncrasies by averaging the use of special education across several grades. Our basic strategy is thus to provide evidence on conditional causal peer effects by exploiting the variation over years in the average proportion of eligible students in a number of grades in each school.

The structure of the paper is as follows: Section 2 provides a short presentation of the relevant theoretical contributions; Section 3 introduces the data, characteristics of the Norwegian school system and the empirical strategy. Section 4 presents results, while Section 5 concludes the paper.

2. Theoretical considerations and existing empirical evidence

From the thin theoretical literature on decision making in schools, we draw attention to two contributions. Motivated by the findings reported in the two most outstanding empirical analyses of peer effects in the seventies (Summers and Wolfe (1977), Henderson et al (1978)), Arnott and Rowse (1987) discuss what allocation of students and resources over classrooms most efficiently achieves various educational objectives. In their model; a social planner maximize students' welfare by being able to manipulate peers' average ability and allocate resources differently. Within this setting, the optimal allocation of students and educational expenditures over classrooms when peer group effects are present is sensitive to the curvature properties of the educational production function. Since the peer group is characterized by average ability; the Arnott and Rowse model is not directly applicable to the problem at hand. In a slightly more general model we could think of optimal resource allocation as being conditional on the student body, the latter being characterized either by ability level or, as in our case, on the level of disruption.

Somewhat more recently, Lazear (2001) treats classroom education as a public good with congestion. When one student disrupts a class, learning is reduced for all the other students. The key variable in Lazear's model is p, which is the probability that a student will behave well. When all students in a class of size n are behaving in the same way, disruption occurs $1-p^n$ of the time. Disruption is decreasing in p and increasing in n. By assuming that students in the same class have different p's, this model can be used to discuss the allocation of special education resources. For principals and school owners, the individual student's p may be observable, and they might act to reduce the consequences from having students with a low p. Taking care of the students with the lowest p by providing special education, instruction time for ordinary students in the classroom, p^n, increases. This increase may come from two sources: 1) more special education may effectively reduce n, that is, by removing or taking care of a student may reduce the group size, the remaining students are exposed to a more favorable teacher-student ratio. 2) Special education to a student with a low p will allow more time for instruction, either by increasing this particular student's p or by taking the student with a low p out of the class, making the expression of p^n larger. Importantly, if school decision makers systematically allocate special education resources to boys, the potential gender peer effects might be neutralized.

The main lesson to take home from the two theories that are given a short presentation here is that school decision makers might be able to manipulate the social interactions in class. We could still talk about causal peer effects, but not about unconditional causal peer effects. It is fair to say that this insight is not at the forefront of the current non-experimental empirical peer literature, implying that there is a gap to fill. This is the motivation for the present paper.

A small number of recent empirical analyses focus on the negative externalities that are associated with distinct subgroups of classmates. Econometrically these studies highlight the problems associated with parental sorting. One important example is Lavy and Schlosser (2010), who use Israeli data to find that an increase in the proportion of girls in the class improves both the boys' and the girls' academic achievements; a 20 percentage point increase in the proportion of female peers increases test scores by approximately 4–5 percent of a standard deviation in the students' test score distributions. Moreover, they provide evidence that the mechanism appears to be that a higher proportion of females in a class leads to a better 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. In Israel, however, learning disabled students are taken out of the school and placed in special education schools. The authors do not provide information on the selection of these students, or the criteria for removing them from ordinary instruction.

Other empirical studies try to pin down the negative effects associated with more narrowly defined "bad apple" peers. Figlio (2003) and Fletcher (2010) find negative effects from attending a class with disruptive students. Figlio (2003) finds that disruptive classmates reduce overall mathematics achievement and increase the likelihood that other classmates will become disruptive. Fletcher (2010) finds that students with classmates that have serious emotional problems, score significantly lower than other students. One contribution using

Norwegian data is Bonesrønning (2006), who highlights the negative externalities related to classmates from dissolved families. All these studies identify student subgroups associated with negative spillovers, but are silent about the learning environment or more specifically, about the use of special education. An implicit assumption is that school actors do not respond to student subgroups that the researchers are able to identify as being challenging for the learning environment.

If school actors systematically try to mitigate potentially negative peer effects by actions that are unobserved or ignored by the researcher, we should expect that studies using non-experimental data provide estimates of peer effects that are biased downwards. (Note however, that if studies using non-experimental data are unable to do away with all self-selection among parents the bias might be in the opposite direction.) We are aware of no experimental studies of bad apples, but experimental studies of ability peer effects tend to report larger peer effects than non-experimental studies of the same phenomenon. Epple and Romano (2011) provide an overview: while studies exploiting randomization in primary education (Duflo, Dupas, and Kremer, 2008; Kang, 2007; and Whitmore, 2006; Graham, 2008) find large peer achievement effects ranging from .20 to .60, studies using non-experimental data and fixed-effects methods generally lead to peer effect estimates that are smaller. Using fixed effects Hanushek, et. al. (2002) find an effect of .15, while Vigdor and Nechyba (2004) and Zabel (2008) obtain estimates on the order of .05. Using selection-on-observables, Ding and Lehrer (2007) conclude that the peer effect coefficient is on the order of 0.08 to 0.15. Hoxby and Weingarth (2006) stand out as an exception among the non-experimental studies, finding peer effects on the order of 0.25 in the linear-in-means model, which is within the range reported by the experimental studies.

We are unaware of existing empirical studies that highlight the potential importance of school actor behavior on peer group effects. However, a few studies investigate the effects of special education resources. Hanushek, Kain and Rivkin (2002) find that special education students do not harm the performance of ordinary students, and that ordinary students that are exposed to eligible students that are not learning disabled, emotionally disturbed or speech impaired will have higher achievement gains than other individuals. A study by Friesen and Krauth (2008) suggests that there are negative spillovers from classmates with special needs (learning and behavioral difficulties) on academic performance.

Thus, our reading of the existing empirical literature (experimental and non-experimental studies) and the theoretical literature indicate that school decision makers might act in ways that dampen potential negative spillovers between students. Causal peer effects might be conditional on such behaviors. The purpose of the below analyses is to examine whether conditional peer effects appears in our data. These data are presented below.

3. Institutions and Data

Institutions

The analyses presented here use data from Norwegian public schools. Ninety-eight percent of the Norwegian elementary school students are enrolled in these schools. Norway has a federal system, where multi-purpose municipalities (about 430) run the public elementary and lower secondary schools (a total of about 2900) subject to national laws and regulations. The municipalities are financed by local taxes – tax rates set by the national government - and national grants allocated on the basis on municipality characteristics. In this multi-purpose, fixed budget setting, the municipal councils face trade-offs between schools, kindergartens, care for the elderly, and some more purposes.

Enrollment in a public school is determined by the location of the students' residence. When the student is six years old, he/she enters the neighborhood school. Parents who want to change schools because of, for instance, a "bad draw in the gender composition lottery" experience high costs. Classes segregated by gender are not allowed, and the school organization is characterized by home classes, implying that the students spend almost all of their classes with the same peers. Early/late starting students and grade retention are extremely rare.

Allocations of special education resources are important in this paper. In Norway, the right to special education is regulated by national law. The law says that students that do not benefit from the ordinary teaching are entitled to special education. Eligibility is determined by experts hired by the municipalities. Eligible students are assigned to one of the following categories: visual or hearing impairment, communication problems, brain damages, learning disabilities, concentration problems, or misbehavior (related to ADHD, other diagnoses or no specific diagnoses). Having received a diagnosis, eligible students are assigned a total number of hours in special education per year. In a survey to the municipalities in 2009, about half of the municipalities answered that this decision was taken at the municipal level, while the other half answered that this decision was decentralized to the schools. In the former case, the municipal officers allocate the total educational budget to the local schools, while also determining the allocation of resources between special education and alternative uses for each school. In the latter case, the schools face a within-year fixed budget, and have to allocate their resources between alternative uses.

The organization of special education is guided by the principle of full inclusion. Thus, most students that are deemed eligible are taught in ordinary classes by adding a special education teacher or an assistant. Alternatively, the special education students are taught in smaller groups of eligible students for a limited number of hours, or sometimes tutored (alone) for a limited number of hours. Only students with the most serious kinds of retardation are taught in special schools. In 2001 2.1 percent of the students in the capital of Oslo and 0.4 percent of the students in the rest of the country were enrolled in special schools.

While the institutions governing special education have stayed unchanged since 1998, the parliament adopted a new governing system for compulsory schooling in the period 2003-2006. The system was changed from an input-oriented system to a system with accountability elements; the important new elements being fewer and clearer goals, decentralization of decision-making to the schools and accountability for teachers and school leaders. The reform was decided in the parliament, with the explicit assumption that the 430 municipalities changed their own local governing systems according to the reform. The latter has happened to a limited extent only; implying substantial variation in local governing systems (see Bonesrønning, 2013). Before the reform period, parents had access to no information about the relative academic performance of their children and schools. National tests were introduced for all students in the 5th and 8th grade in 2004, but were withdrawn due to resistance from the teachers' union and some other stakeholders (and these data are no longer available). The tests were reintroduced in 2007, changing the informational environment fundamentally. Before the introduction of national tests, no information existed about the performance of the 5th graders. Afterwards, better informed parents might have increased their demand for special education, and the teachers' incentives to ask for diagnoses may have been strengthened. The municipalities might have been more or less prepared to handle this situation.

Data

We use data from three adjacent cohorts (2007-2009) of 5th graders in the Norwegian elementary school. All 5th grade students (except exempted students) are tested in the early autumn in math, reading in Norwegian and reading in English. We do not include 8th graders. The reason is that the transition from the elementary to the lower secondary school takes place between the 7th and 8th grade, implying that most students are exposed to new peers from the 8th grade.

We combine information about the 5th graders' test results in 2007-2009 with data from the Compulsory School Information System (GSI). GSI is a school administrative system that collects information from all the school principals in elementary and lower secondary schools in Norway. From 2006 and onwards, GSI reports the amounts of special education resources by grade-in-school. In addition, teaching hours, and the numbers of students, assistants, administrators and so forth are reported. Statistics Norway has supplemented these data with information about individual students (gender, birth order, ethnicity), their parents (education, income), the families (size, structure), and municipality identifiers.

The outcome variable is constructed by standardizing and summarizing the individual scores from national tests in all three subjects. The sum score is standardized to a mean of zero and a standard deviation of one. These aggregation practices potentially conceal important differences in peer effects across the subjects. We have performed analyses for each subject, but it turns out that these analyses don't provide any important additional insights. Therefore we report results using aggregated outcome measures only.

The two explanatory variables of main interest are the proportion of boys in the grade and the proportion of students receiving special education. Approximately half of the population is boys, but both the between-schools and the within-grade-in school variation in the gender composition are relatively high. When decomposing the standard errors we find a within-grade standard deviation of 0.08, compared to a total standard deviation of 0.12, indicating sufficient variation in both dimensions. In the analyses we use the proportion of boys measured at the grade level, not the classroom level. The reason is that we have no information about student allocation across classes within schools. These practices might have been preferable also if information about the classroom composition had been available - due to the potential endogeneity of the classroom peer composition.

The level of special education in the school is measured by the proportion of students that are deemed eligible to special education. This is not an obvious choice. We have considered alternative measures, such as the hours of special education per eligible student or hours of special education per student in the class, but have decided to use the proportion of eligible students – the main reason being that this is the only measure that is reported at the grade level. The reader should be aware that within each school year, the school principals face a trade-off between the number of eligible students and the number of special education hours per eligible student. Over the three year period considered here, the school budgets have increased, implying that the principals to some extent have been able to expand the number of eligible students without decreasing the number of hours per eligible student. A priori it is not obvious whether the alternative "a small student group and many hours of treatment" or the alternative "a large student group and fewer hours of treatment" has the largest impact on the gender peer effects and the learning environment. (In an appendix we discuss whether other school inputs are time-varying in the period 2007-2009)

In the analyses we (initially) use the proportion of 5th grade students that are deemed eligible to special education in the 4th grade to characterize the learning environment. The national tests are taken early in the 5th grade, which implies a very short period of treatment in the 5th grade compared to the year-long treatment in the 4th grade. In addition, by using the level of special education in the year prior to the 5th grade, we overcome problems with schools that have increased their level of special education after the students are tested. Approximately 4.5 percent of the 5th graders received special education in the 4th grade in 2007, increasing to 5.0 percent in 2008 and 5.8 percent in 2009, implying an increase in the proportion of eligible students of almost 30 percent. Table 1 provides descriptive statistics for the two explanatory variables of main interest. Additional descriptive statistics are provided in Appendix Table A1.

[Table 1 about here]

More about patterns in the data

It is useful to look more carefully into the characteristics of the 2007-2009 expansion in the proportion of special education students. First, the proportion of eligible students is higher in the high grades: for example in 2007, 4.3 percent of the 1st graders, and 10.6 percent of the 10th graders, received special education. Second, the across-grade differences are reduced in the three year period: the increase in the proportion of eligible students in 1st – 4th grades is about twice the increase in 5th -7th grade, probably reflecting the government's emphasis on early intervention. The latter characteristic – that the environment in which gender peer effects occurs has changed relatively more in the early grades – is potentially important for the analyses.

Some existing evidence (i.e., Lavy and Schlosser (2010)) indicates that gender peer effects are larger in environments characterized by low-SES families, the most likely mechanism being that boys from low-SES families are more likely to misbehave in class. Motivated by these findings we have investigated whether the increase in the proportion of eligible students differs across schools according to the average education of the 5th grade parents. We find that all types of schools have experienced an increase in the proportion of eligible students in the period 2007-2009, but that schools with average parental education 0-1 standard deviations above the population mean have experienced a larger increase in eligibility than schools with average parental education 0-1 standard deviations below the mean. Thus, the environment has changed most in schools where we would expect the gender peer effects to be relatively small. Figure 1 illustrates.

[Figure 1 about here]

We have investigated the relationship between the two explanatory variables of main interest by regressing the proportion of eligible students in the 4th grade against the proportion of boys, together with year dummies, while controlling for other student body characteristics.

[Table 2 about here]

The regression reported in Table 2, column 1, includes the proportion of boys and year dummies (2007 is the year of reference) as the only explanatory variables. Control variables are added in columns 2-4, and column 5 report the results from a specification with school fixed effects. Reading across the columns, it is evident that an increase in the proportion of boys is associated with a statistically significant increase in the proportion of special education students, and moreover, the proportion of eligible students has increased substantially in the three year period. Controlling for the student body composition, the proportion of eligible students is 1.1 percentage points higher in 2009 than in 2007. These results underline the arguments made in the present paper: in this three-year period the school decision makers seem to have responded to the student body composition in ways that might dampen the negative externalities that are associated with boys. Moreover, the statistical significant and substantial increase in the proportion of eligible students through the short time period provides an opportunity for investigating the hypothesis that the size of the gender peer effects depends on the amount of special education resources.

Finally, before turning to the more rigid analyses, we estimate education production functions (EPF) separately for 2007, 2008 and 2009 to illustrate that the level of special education *might* matter for the size of the gender peer effects. Table 3 reports the results from this exercise.

[Table 3 about here]

Two specifications of the EPF are provided for each year; the proportion of eligible students being included as a control variable in the second, but not in the first, of these. The point estimate for the proportion of boys is about 0.16 in 2007 (a slightly smaller effect than reported in Lavy and Schlosser (2010)), and decreases to a quarter of this size in 2009. Thus, an increase in the proportion of eligible students goes together with a decrease in the gender peer effect. There are at least two reasons why this is not decisive evidence for context-specific peer effects. First, the gender peer effects are derived from the between-grade-in-school variation in the proportion of boys. Second, no explicit relationship between the gender peer effects and the school context is established.

3.2 Empirical Strategy

To deal with the potential endogeneity of the peer group that originates from parental sorting we follow the established practices and exploit variations in the proportion of boys over time within the same school.

Using repeated cross-sectional data (three adjacent cohorts), we estimate equation (1) below:

$$(1)\quad A_{ist} = \beta_s + \beta_1 P_{st} + \beta_2 F_{ist} + \beta_3 X_{st} + u_{ist}$$

where i denotes the individuals, s denotes the schools and t denotes time. A_{ist} is the achievement measure of student i in school s and year t. β_s is a school effect and P_{st} is the proportion of male students in school s at time t. F_{ist} includes characteristics of the individual student and his/her family. This vector contains information about gender, immigrant status, parental education, parental income, family type and birth order. X_{st} is a vector of the school characteristics in school s at time t, and contains information about time-varying school characteristics such as cohort size, the number of assistants per student, the proportion of teachers without approved education, the proportion of male teachers and the level of special education; these characteristics vary over time within the same school. The coefficient of interest is β_1, which captures the effect on achievement of having more male peers in the same grade.

Above we have discussed that school authorities might be able to respond in a way that mitigates peer effects. Equation (1) takes into consideration that the peer group composition is correlated with special education resources, but not that the peer group effects might be conditional on the special education resources. We address this issue by including an interaction between the proportion of boys and the proportion of students with special education, as portrayed in equation (2).

$$(2)\ A_{ist} = \beta_s + \beta_1 P_{st} + \beta_2 F_{ist} + \beta_3 X_{st} + \beta_4 SE_{st} + \beta_5 SE_{st} P_{st} + \mu_{ist}$$

This equation differs from equation (1) in two respects. The time-varying proportion of eligible students (SE_{st}) is taken out of the X-vector, and an interaction term between the peer variable and the proportion of eligible students is included. If the point estimates β_1, β_3 and β_5 are statistical significant, this will be a first direct indication that the estimated peer effects depend on the use of special education. Within this framework we also report analyses where we exclude from the analyses schools that always use special education and schools that never use special education, and thus highlight the importance of the learning environment for the size of the peer group effects by exploiting schools that go from special education one year to no special education the next year.

This fixed effects approach takes away some part of the endogeneity of special education resources by making use of the variation in eligibility across adjacent cohorts within schools. However, it remains that we neither observe the student behavior that qualify for special education nor the teacher characteristics that influence the likelihood that a student misbehaves in class. We would like to get rid of the variation in eligibility that is due to idiosyncrasies in the (adjacent) cohorts for whom we observe academic performance. To achieve this, we make use of two different approaches. First, we characterize the special education environment for the 5th graders by using the average proportion of eligible students in the 1th-4th grade for each of the years 2007-2009. This is a simple procedure to cultivate the (school-specific) time trend in eligibility by "aggregating away" much variation that is due to student and teacher idiosyncrasies across grades. Second, we separate the municipalities by their proportion of public employees. This is motivated by prior investigations (Bonesrønning, Iversen, and Pettersen 2012) showing that the proportion of eligible students increases more in municipalities with large proportions of public employees. In line with Iversen (2013) we argue that this reflect that such municipalities have not installed robust governing systems to handle the increasing demand for diagnoses. Thus, we exploit that the variation in proportions of eligibility across municipalities with different proportions of public employees over time is due to factors that are unrelated to student and teacher idiosyncrasies.

4. Results

The result section consists of two parts. First, we follow the existing empirical literature (i.e. Lavy and Schlosser (2010), Hoxby (2000)), and estimate gender peer effects taking account of parental sorting. In the next part we estimate gender peer effects taking account of parental sorting and the actions of school decision makers, following the theoretical implications and the discussion in Mofitt (2001). Based on the theoretical considerations above, and the discussion of experimental and non-experimental studies on peer effects, our assumption is that school authorities observe positive and negative externalities in the classrooms, and manipulate them in order to achieve different educational objectives. We expect that the observed gender peer effects will be smaller with intervention than without intervention.

4.1 Estimating gender peer effects – taking account of parental sorting

For comparison, the two first columns of Table 4 report the results from estimating equation (1) using OLS. The three last columns report results from school fixed effects estimations. The point estimates for the proportion of boys are negative throughout the table. The estimates drop somewhat, and lose some statistical significance, in the school fixed effects specifications compared to the specifications that make use of both within- and between-grade-in-school variation. Nonetheless, the point estimates are statistically significant in the first four columns, and the estimated size of the peer effect is about the average of the peer effects estimated for each of the three years (see Table 3). Based on the largest point estimate (column 1), an increase in the proportion of boys from 40% to 60% will affect achievement by roughly 2% of a standard deviation; which is a little less than half the effect reported by Lavy and Schlosser (2010).

[Table 4 about here]

The estimates reported in the two right hand columns are estimated off what is basically idiosyncratic variation in the gender composition of subsequent cohorts in the 5th grade. The claim that the effects are identified off idiosyncratic variation in the peer composition is reinforced by a conventional balancing test where family characteristics are regressed against the proportion of boys, finding no evidence that the peer measure is correlated with any of these measures. The latter results are reported in Appendix Table 3. The applied method is thus close to the by now standard approach to estimating causal gender peer effects. However, in the first of the fixed effects-specifications we have suppressed that special education is time-variant. In the second specification, the proportion of eligible students is included among the independent variables. When the proportion of eligible students is included, the estimated gender peer effect drops somewhat in magnitude, and also, it is no longer statistically significant at conventional levels. The latter specification is also open to criticism. First, according to arguments made earlier, the peer effects are conditional on the learning environment. The specification presented here does not address this issue. Second, the proportion of eligible students is an endogenous variable, and as such, a bad control. We start with the first of these objections, that is, we report results from estimating equation (2).

4.2 Estimating gender peer effects - taking account of school and municipal decision making

This section addresses our main purpose which is to evaluate the potential conditionality of gender peer effects. We use different strategies. Column 1 in Table 5 provides results from estimating equation (2) with school fixed effects. The point estimates for the proportion of boys, the proportion of eligible students, and their interaction are all statistical significant. The signs for the proportion of boys and the interaction variable are negative and positive, respectively, indicating that a large proportion of eligible students are associated with a smaller gender peer effect. If no students are deemed eligible the gender peer estimate is equal to -0.11, and if 15 percent of the students are deemed eligible, a change in the proportion of boys is associated with no changes in student performance. These findings might indicate that causal gender peer effects are conditional on the use of special education, but since the variation in special education across adjacent cohorts within schools reflects decisions made by school actors we do not make any strong claims about conditionality at this point.

[Table 5 about here]

It might be useful to speculate about the direction of the bias in the estimates provided above. For this purpose we have chosen a subsample of schools that do not allocate any special education resources for one or two of the three adjacent cohorts of 5th graders we are investigating. The decision to provide special education to a cohort potentially reflects many factors, student and teacher characteristics being among the most important. No special education is likely to be provided to cohorts characterized by a combination of small idiosyncratic problems among the students and high teacher quality. It seems unlikely that no special education resources are allocated to the cohorts with the largest idiosyncratic problems. We therefore expect the gender peer effects to be "relatively small" when no special education resources are used. Here “relatively small” have (at least) two interpretations. First, the estimates for gender peer effects based on cohorts in schools that have no special education are likely to be smaller than the peer effects we would have estimated if no schools actually used special education. Second, the estimates for gender peer effects based on schools that have no special education are likely to be smaller than the peer estimates based on schools that use special education if special education has modest dampening effects on negative externalities, but larger if special education effectively dampen negative externalities.

The results from the regression analyses for this subsample are reported in Table 5, columns 2 and 3. Note that since school enrollment and the grade enrollment appear to be larger in the years when special education resources are used, we control for this in the regressions. For cohorts-in-schools that have no special education students, we estimate significant negative gender peer effects. A 20 percentage point increase in the proportion of boys decreases student achievement by a little more than 3 % of a standard deviation, which is a larger effect than reported above for the entire population. For cohorts-in-schools that use special education resources, we find no significant gender peer effects.

Since the gender peer effects that are present when no special education resources are used, are estimated from cohorts that we conjecture have favorable unobservable characteristics, these gender peer effects are probably smaller than we would find for the entire population in the absence of special education. Moreover, the absence of significant peer effects when special education resources are used *might* indicate that these resources effectively remove negative externalities. Note that these arguments originate from analyses of within-school across-cohort variation in gender composition. That is, we have not taken decisions made by school principals and municipality officers into consideration.

We would like to provide unbiased estimates showing how an increase in special education – that is not due to unobservable teacher and school leader characteristics – affect the gender peer effects. For this purpose we have exploited the across-the-board increase in eligibility across the three cohorts under inspection. As stated above, the Norwegian elementary school has experienced a rapid increase in the proportion of eligible students, from 4.5 percent of 4th graders in 2007 to 5.8 percent in 2009. Similar increases have taken place for all other grades. We substitute the proportion of eligible students in the 4th grade with the average proportion of eligible students in the 1st -4th grades. Using the average proportion of special education students from 1st -4th grades is a simple procedure to cultivate the (school specific) time trend in eligibility by "aggregating away" much variation that is due to student and teacher idiosyncrasies across grades.

The results from these analyses are reported in Table 6. From column (1) it is evident that the estimate for the proportion of boys is of the same size as in Table 4. However, the estimate for the interaction between the proportions of boys and eligible students is smaller and less precisely estimated than there, indicating that the within-school variation in the average proportion of eligible students in grades 1-4 is too small to generate precise estimates. The sign of the relationship between special education and the peer effect is however the same. This exercise then provides some limited support for the hypothesis of conditional peer effects.

[Table 6 here]

The analyses so far have all given indicative evidence that gender peer effects are conditional upon the level of special education. Even when substituting the use of special education in 4th grade by the level on 1st-4th grad we see this relationship, although not significant. However, these analyses have not addressed potential heterogeneity among the municipalities. There are reasons to believe that this is an issue, notably because the public schools are embedded in a federal structure where the municipalities very much decide their own governing systems. A national reform in 2006, called “Knowledge Promotion”, encouraged municipalities to change their governing systems and to include more accountability elements. Bonesrønning (2013) provides evidence that the national accountability reform is poorly implemented in municipalities with high proportions of public employees, and Iversen (2013) shows that municipalities that have not implemented the national reform experience higher growth in the proportion of students that are eligible to special education in the period 2007-2009.

Iversen (2013) argues that municipalities with accountability systems are better equipped to keep costs and special education at a lower level (for instance, the informational asymmetries within schools might be less severe in accountability systems, implying that school principals are in a position to push harder for customized training). For the same reasons these municipalities will be in a better position to dampen boys’ misbehaviors by applying other means than special education resources. If this is the case, we might find smaller gender peer effects in the absence of special education resources in these municipalities, and consequently, we might find that special education resources will have smaller effects on the learning environment. In short, reformed municipalities might have other tools than special education resources to handle boys’ misbehavior.

Motivated by such conjectures, we have separated the population of municipalities into two subgroups, one with less – and one with more - than the average population proportion of public employees (which seems like a variable that is fairly exogenous to student achievement). We have estimated the relevant equations for each of these subsamples, implicitly treating the proportion of public employees as an exogenous municipality characteristic.

Panel A in Table 7 reports the results from estimating equation (2), using the observed proportion of special education students in the 4th grade as the measure of special education. Two results are important. First, we find that for municipalities with less than average proportions of public employees there are no gender peer effects. As discussed above, these municipalities will be in a better position to dampen boys' misbehavior by applying other means than special education resources. Second, for municipalities with more than average proportions of public employees the gender peer effects depend on the proportion of special education students: quite large effects (50% larger than reported above) for schools that use no or only small amounts of special education, and no effects for schools that have above 20 percent of special education students.

[Table 7 here]

Panel B in Table 7 reports results from regression analyses where we have substituted the average proportion of special education students in the 1st -4th grades for the observed proportion of 4th grade special education students. The findings for municipalities with small proportions of public employees are basically unchanged. For the subsample of municipalities with high levels of public employees, the estimated coefficient for the interaction variable is now larger in absolute value, but less precisely estimated (no longer statistically significant), than in the upper panel. This is in line with the analyses above when using the special education level at 1st-4th grade. There is less variation (more aggregated), such that even if the coefficients are similar in magnitude, the effects will be less precisely estimated.

Based on the results in Table 7, it seems evident that gender peer effects do not appear everywhere across the universe of municipalities. On average, and independent on the use of special education resources, we find no gender peer effects in municipalities with small levels of public employees. In municipalities with high proportions of public employees we find significant gender peer effects; the largest gender peer effects occurring among the municipalities that have the smallest proportions of special education students (although not all schools seem to experience this effect). We do not claim that these findings provide decisive evidence in favor of the hypotheses we have presented to motivate the separation of municipalities into two subgroups. Much remains to be understood about why the gender peer effects seem to vary across different types of municipalities.

Nonetheless, we think that the exercises presented in this section provide indicative evidence that gender peer effects are context-specific. In this respect we have throughout the paper emphasized the importance of special education resources, but the analyses presented right above indicate that also institutional characteristics might be important.

4.3 Robustness checks

There are other concerns regarding the empirical specification presented above. First, the relationship between the two variables of main interest could result from non-linearity in the gender peer effects. Because the boys dominate the special education group, the special education term could capture the non-linearity, and we would, in fact, be estimating a more flexible gender peer effect. We have estimated a model where the special education term is excluded, and a squared term of the gender peer term is included. This model, which is not reported in the tables, does not support the hypothesis that the peer effects are non-linear.

Another concern is that some special education students are included in the analyses. In fact, most of these students are excluded. Principals have the opportunity to exclude all students with special education from the testing pool. Nonetheless, Table A2 present results where we have excluded the students with the highest probability of receiving special education. The results are robust for these specifications.

A falsification check is performed by using the proportion of boys in the 4th grade instead of the proportion of boys in the 5th grade as explanatory variable. The proportion of boys in the 4th grade should not have any effect on achievement for the 5th graders. These results are reported in Table A3 and show that the proportion of boys in the 4th grade has no effect on the achievement of the 5th graders.

Inspired by Altonji, Elder and Tabor (2005), we also test the robustness of our peer estimates by progressively adding different types of observable student covariates to the regression models. As illustrated in Table 3, adding individual or family covariates does not significantly change the coefficient. Following the intuition in Altonji et al., this result indicates that the potential bias from unobservables is small.

We also construct a peer proxy that corrects for the individual students' sex. In our earlier analyses, the individual student's sex contributes to the gender composition. This contribution could create a small bias in the estimates as well. However, by correcting for the individual students' sex, the estimates and their statistical significance hardly changes. This evidence is also provided in Table A3.

5. Conclusions

Much of the empirical literature that focuses on peer group effects in schools and classrooms does not explicitly address the school actors' actions. More generally, the potential importance of the context in which the interactions occur, is not much highlighted in the empirical literature. On the other hand, theoretical contributions to the peer effects literature highlight the role of decision making within schools. If educational planners act in order to maximize different educational outcomes, any observed negative peer effects will be smaller than a situation without educational planners' actions. Since the most common non-experimental strategies of estimating peer effects do not (fully) take such behaviors into account, this might be the reason why non-experimental studies tend to report smaller effects than effects found in experiments.

In the present paper we investigate how gender peer effects are affected by the use of special education in the Norwegian elementary school. Important features of the Norwegian context are that special education is regulated by national law saying that students that do not benefit from ordinary teaching are entitled to special education, that student misbehavior - related to ADHD, other diagnoses or no specific diagnoses - is one of the categories qualifying for special education, and that special education is provided within the classroom.

We use data for the period 2007-2009, basically exploiting a rapid increase in the proportion of students receiving special education in this period to provide credible, but still indicative, evidence that the use of special education resources work to dampen gender peer effects. The evidence is indicative because the inherent endogeneity problems are severe, notably it is hard to deal with the endogeneity of special education. Nonetheless, we think that the present paper – by highlighting that school actors act to dampen potential negative externalities in the classroom – contributes to the current empirical literature that investigate peer group effects within schools.

Acknowledgements

We thank Lars Erik Borge at NTNU and Colin Green at Lancaster University for very useful comments. Bjarne Strøm, Torberg Falch, Marianne Haraldsvik, Arnt Ove Hopland, Kaja Høiseth Brugård and Rune Reiling at NTNU are gratefully acknowledged for their comments on an early draft. Iversen would also like to thank the seminar participants at the Norwegian Economic Association (Bergen) in 2011, the participants of the workshop in Trondheim: “Educational Systems and Human Capital” in 2011, and the participants at the annual EALE conference in Bonn 2012. This research is funded by the Norwegian research council and is part of a larger project on governance, management and performance in the Norwegian educational system. Any errors are our own.

References

Altonji, J. G., Elder T.E., and Tabor C.R. (2005) Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, 113: 151-184

Angrist, J. and Lavy, V. (1999). Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement. *The Quarterly Journal of Economics 114(2) 533-575.*

Black, S.E., Devereux, P.J., and Salvanes K. (2010). Under Pressure? The Effects of Peers on Outcomes for Young Adults. NBER *Working Paper No. 16004*

Bonesrønning, H. (2008) Peer Group effects in education production: Is it about congestion? *The Journal of Socio-Economics 37 (1) 328-342.*

Bonesrønning, H and J.M.V Iversen (2008) Suksessfaktorer i grunnskolen: Analyse av nasjonale prøver 2007. *SØF Rapport 5/08.*

Bonesrønning, H. and Iversen J.M.V. (2013) Disadvantaged students in the early grades: Will smaller classes help them? *Education Economics*, 21(4), 305-324 .

Bonesrønning, H., Iversen, J.M.V., and Pettersen I. (2010). Kommunal skolepolitikk etter kunnskapsløftet. *SØF Rapport 07/10.*

Bonesrønning, H. and Haraldsvik, M (2012) Peer effects on student achievement: Does the educational level of your classmates' parents matter? *Unpublished manuscript*

Bifulco, R., Fletcher J.M., and Ross, S.L. (2011) The effect of Classmate Characteristics on Individual Outcomes: Evidence from the Add Health. *American Economic Journal: Economic Policy.* 3(1), 25-53.

Ding, Weili and Stephen Lehrer (2007) "Do Peers Affect Student Achievement in China's Secondary Schools?" *Review of Economics and Statistics*, 89(2), pp. 300-312.

Duflo, Esther, Pascaline Dupas, Michael Kremer (2011) "Peer Effects and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." 101(5), 1739-74

Figlio, D.N. (2007). Boys Named Sue: Disruptive Children and Their Peers. Education Finance and Policy 2(4), 376-94.

Fletcher, J. (2010). Spillover Effects of Inclusion of Classmates with Emotional Problems on Test Scores. *Journal of Policy analyses and Management (27(1) 69-83.*

Friesen, J., Hickey, R., and Krauth B. (2010) Disabled Peers and Academic Achievement. Educ*ation Finance and Policy 5(3), 317-348.*

Hanushek, E. A., Kain, J.F., and Rivkin S.G. (2002). Inferring program effects for special populations: Does special education raise academic achievement for students with disabilities? *Review of Economics and Statistics, 84(4).584-599*

Hanushek, Eric A., John F. Kain, Jacob M. Markman, and Steven G. Rivkin, "Does Peer Ability Affect Student Achievement?" *Journal of Applied Econometrics*, October 2002, pp. 1-26.

Henderson, J. Vernon, Peter Mieszkowski, and Y. Sauvageau, "Peer group effects and educational production functions," *Journal of Public Economics* 10 (1978) 97-106.

Hoxby, Caroline M. and Gretchen Weingarth, "Taking Race out of the Equation : School Reassignment and the Structure of Peer Effects,"

Hoxby, C. (2000). Peer effects in the Classroom: Learning from Gender- and Race Variations. *NBER Working Paper No. 7867.*

Jahnsen, H., Nergaard, S.E., and Flaatten S.V. (2006) I randsonen: Forekomst og Organisering av smågruppetiltak for elever på ungdomstrinnet som viser problematferd og lav skolemotivasjon. *Utdanningsdirektoratet 2005.*

Kang, Changhui, "Classroom Peer Effects and Academic Achievement: Quasi-Randomization Evidence from South Korea," *Journal of Urban Economics,* 61 (2007) 458 495.

Lavy, V and A. Sclosser (2011) Mechanisms and Impacts of Gender Peer Effects at School. *American Economic Journal: Applied Economics,* 3 (2), 1-33

Lazear, E. P. (2001). Educational Production, *The Quarterly Journal of Economics 110(4), 777-803-693.*

Leuven, E., Oosterbeek H., and Rønning M. (2008) Quasi-experimental estimates of the effect of class size on achievement in Norway. *Scandinavian Journal of Economics 110(4), 663-693.*

Manski, C. (1993). "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies, Vol. 60, No. 3, 1993, pp. 531-542*.

Mckee, G., Rivkin, S.G., and Sims K.R.E. (2010). Disruption, Achievement and the Heterogeneous Benefits of Smaller Classes. *NBER Working Paper No. 15812.*

Moffitt, Robert, "Policy Interventions, Low-Level Equilibria, and Social Interactions,´ Social Dynamics, Blume and Durlauf (eds), pp. 45-82, Washington, DC: Brookings Institution Press.

Nordahl, T. and Hausstätter R.S.(2009). Spesialundervisningens forutsetninger, innsatser og resultater. Høgskolen i Hedmark.

Summers, Anita and Barbara Wolfe, "Do schools make a difference?" American Economic Review 67, (1977) 639-652.

Strøm, B., Borge, L.E., and Haugsbakken H.(2009). Tidsbruk og organisering i grunnskolen: Sluttrapport. SØF-RAPPORT no. 04/09.

Utdanningsdirektoratet (2009:) Veileder til opplæringsloven om spesialpedagogisk hjelp og spesialundervisning.

Vigdor, Jacob and Thomas Nechyba, "Peer Effects in Elementary School: Learning from "Apparent" Random Assignment," Working Paper, Duke University and NBER, November 22, 2004.

Vigdor, Jacob L. and Thomas S. Nechyba (2007). "Peer Effects in North Carolina Public Schools" in Schools and the Equal Opportunity Problem, P.E. Peterson and L. Wößmann, eds. MIT Press.

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.

Zabel, Jeffrey (2008). The Impact of Peer Effects on Student Outcomes in New York City Public Schools," *Education Finance and Policy*, 3(2),197-249

Tables and figures

Table 1: Descriptive statistics for the proportions of boys and students eligible to special education. By year.

	2007	**2008**	**2009**
Proportion of boys	0.51	0.51	0.51
	(0.17)	(0.17)	(0.16)
Proportion of eligible students	0.053	0.058	0.065
	(0.078)	(0.081)	(0.086)

Table 2: The relationship between the proportion of eligible students and the proportion of boys.

	Proportion of eligible students				
	OLS	OLS	OLS	OLS	FE
Proportion of boys	0.038***	0.039***	0.039***	0.037***	0.033***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.007
Proportion of students from intact families			-0.018***	-0.021***	-0.019**
			(0.002)	(0.002)	(0.009)
Father's education – grade average		-0.014***	-0.0088***	-0.0044***	-0.0019
		(0.0002)	(0.0003)	(0.0003)	(0.003)
Father's income-grade average			-2.57e-08***	-1.80e-08***	-2.57e-09
			(1.13e-09)	(1.09e-09)	(9.03e-09)
2008	0.003***	0.004***	0.004***	0.002***	0.004***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.001)
2009	0.011***	0.012***	0.012***	0.010***	0.011***
	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.002)
School fixed effects	No	No	No	No	Yes
Constant	0.027***	0.086***	0.090***	0.069***	0.057***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.013)
Observations	145,547	145,547	145,547	145,547	145,547
R-squared	0.013	0.039	0.045	0.085	0.020
Number of schools					2,428

Note: Additional controls not reported include immigration status, family structure, number of brothers and sisters, birth order, grade enrollment, school type, and the proportion of male teachers. Robust standard errors. * significant at 10 % level. ** significant at 5 % level. *** significant at 1 % level.

Table 3: Gender peer effects by year in the period 2007 – 2009, including controls for the proportion of eligible students

	Standardized values of national tests 2007-2009					
	2007	2007	2008	2008	2009	2009
Proportion of boys	- 0.165***	- 0.155***	- 0.101**	- 0.0913**	- 0.0416*	- 0.0405*
	(0.0420)	(0.0421)	(0.0419)	(0.0420)	(0.0238)	(0.0237)
Proportion of eligible students		-0.390***		-0.353***		-0.435***
		(0.0885)		(0.0864)		(0.0807)
Constant	-1.455***	-1.423***	-1.425***	-1.392***	-1.383***	-1.343***
	(0.0424)	(0.0430)	(0.0429)	(0.0436)	(0.0397)	(0.0404)
Observations	49,901	49,901	47,979	47,979	47,521	47,521
R-squared	0.133	0.133	0.133	0.133	0.128	0.128

Note: Robust standard errors clustered at the school level are reported in parentheses. Controls include student gender, immigration status, parental education, parental income, family structure, number of brothers and sisters, birth order, grade enrollment, school type, proportion of male teachers, level of special education and number of assistants. Columns 1 and 2 present OLS results. Columns 3-5 include school fixed effects * significant at 10 % level. ** significant at 5 % level. *** significant at 1 % level.

Table 4: Gender peer effects for the entire population of students in the cohorts 2007-2009.

Standardized values of national tests 2007-2009					
	OLS	OLS	FE	FE	FE
Proportion of boys	-0.103***	-0.104***	-0.0802*	-0.0828*	-0.0689
	(0.0257)	(0.0243)	(0.0453)	(0.0451)	(0.0451)
Proportion of eligible students					-0.407***
					(0.0888)
Individual characteristics	x	x	x	x	x
Control for family background		x	x	x	x
School fixed effects			x	x	x
Time-varying school characteristics				x	x
Control for special education					x
Constant	0.0948***	-0.979***	-0.898***	-0.902***	-0.884***
	(0.0141)	(0.0168)	(0.0262)	(0.0873)	(0.0871)
Observations	145,547	145,410	145,410	145,401	145,401
R-**squared**	0.004	0.126	0.096	0.096	0.097
Number of schools			2,428	2,424	2,424

Note: Robust standard errors clustered at the school level are reported in parentheses. Controls include student gender, immigration status, parental education, parental income, family structure, number of brothers and sisters, birth order, grade enrollment, school type, proportion of male teachers, level of special education and number of assistants. Columns 1 and 2 present OLS results. Columns 3-5 include school fixed effects * significant at 10 % level. ** significant at 5 % level. *** significant at 1 % level.

Table 5: Estimating gender peer effects taking account of the proportion of eligible students

	Standardized values of national tests 2007-2009		
		Reduced sample – schools with no special education one or two years	
	Full sample	**Cohorts with special education students**	**Cohorts with no special education students**
Proportion of boys	-0.110**	-0.0507	-0.151***
	(0.0531)	(0.0484)	(0.0472)
Proportion of eligible students	-0.793***	-0.271***	0
	(0.255)	(0.0991)	(0)
Interaction between proportions of boys and eligible students	0.721*		
	(0.433)		
Constant	-0.862***	-1.366***	-1.211***
	(0.0888)	(0.0508)	(0.0526)
Observations	145,401	35,775	28,110
R-squared	0.097	0.138	0.127
Number of schools	2,424		

Note: Robust standard errors clustered at the school level are reported in the parentheses. Controls include student gender, immigration status, parental education, parental income, family structure, number of brothers and sisters, birth order, grade enrollment, school type, proportion of male teachers, level of special education and number of assistants. All columns reports results on school fixed effects models. * significant at 10 % level. ** significant at 5 % level. *** significant at 1 % level.

Table 6: Estimation results equation (2). Using the average proportion of eligible students in 1st - 4th grades as a substitute for special education in the 4th grade. OLS and FE.

Standardized values of national tests 2007-2009		
	OLS	**FE**
Proportion of boys	-0.112***	-0.110*
	(0.0358)	(0.0584)
Interaction between the proportions of boys and the average proportion of eligible students 1st -4th grades	0.175	0.480
	(0.559)	(0.839)
Average proportion of eligible students 1st – 4th grades	-0.619**	-0.904**
	(0.303)	(0.451)
Constant	-1.283***	-1.089***
	(0.0289)	(0.0587)
Observations	145,401	145,401
R-squared	0.131	
Number of schools		2,424

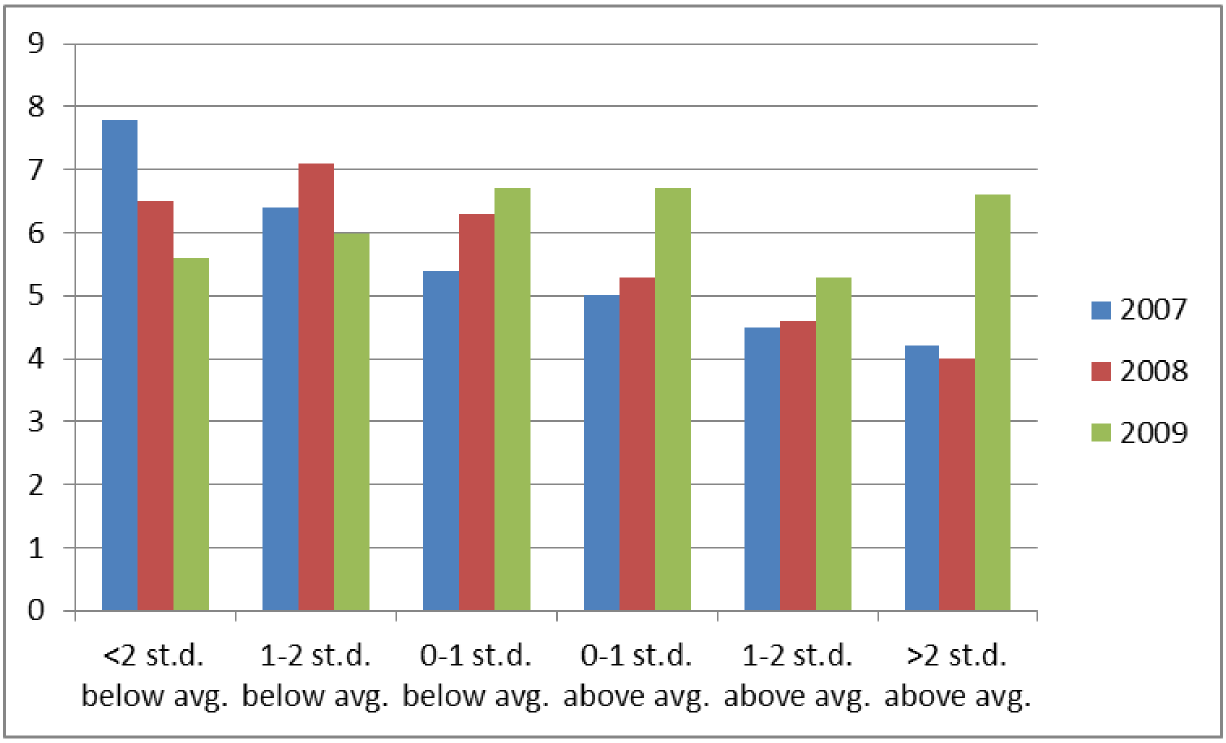
Note: Robust standard errors clustered at the school level are reported in the parentheses. Controls include student gender, immigration status, parental education, parental income, family structure, number of brothers and sisters, birth order, grade enrollment, school type, proportion of male teachers, level of special education and number of assistants. All columns reports results on school fixed effects models. * significant at 10 % level. ** significant at 5 % level. *** significant at 1 % level.

Table 7: Separate estimations of equation (2) for municipalities with low and high proportions of public employees respectively.

Standardized values of national tests 2007-2009		
	Low proportion of public employees	**High proportion of public employees**
Panel A		
Proportion of boys	0.131	-0.174***
	(0.107)	(0.0524)
Proportion of eligible students	-1.056	0.841**
	(1.481)	(0.387)
Interaction between proportions of boys and eligible students	-0.00680	-0.818***
	(0.806)	(0.232)
Constant	-0.902***	-0.848***
	(0.0692)	(0.0415)
Panel B		
Proportion of boys	0.113	-0.178***
	(0.131)	(0.0641)
Interaction between the proportions of boys and the average proportion of eligible students 1st -4th grades	-1.108	0.957
	(2.646)	(0.839)
Average proportion of eligible students 1st - 4th grades	0.129	-1.158**
	(1.368)	(0.472)
Constant	-0.906***	-0.836***
	(0.0803)	(0.0462)
Observations	58,819	86,582
Number of schools	621	1,803

Note: Robust standard errors clustered at the school level are reported in the parentheses. Controls include student gender, immigration status, parental education, parental income, family structure, number of brothers and sisters, birth order, grade enrollment, school type, proportion of male teachers, level of special education and number of assistants. All columns reports results on school fixed effects models. * significant at 10 % level. ** significant at 5 % level. *** significant at 1 % level.

Figure 1: The growth in the proportion of eligible students 2007-2009 across subgroups of students.



Appendix

Table A1: Descriptive statistics for the groups of schools with and without special education

	All students		Reduced sample – schools with no special education one or two years	
	No special education	**With special education**	**No special education**	**With special education**
Mother's education	4.47	4.54	4.5	4.53
Father's education	4.31	4.42	4.34	4.38
Intact families	75.3%	73.8%	75	74.4%
Mother's earnings	263 000	271 000	267 000	269 000
Father's earnings	485 000	494 000	492 000	492 000
First generation immigrants	0.8%	1.1%	0.9%	1%
Second generation immigrants	2.5%	4.1%	2.8%	3%
Proportion of boys	50%	52%	50%	52.3%
Assistants	14.8	16.1	15.1	15.9
Uncertified teachers	7.2%	6.6%	7.1%	6.6%
Male teachers	25.1%	24.4%	24.9%	24.9%
Enrollment	140	219	154	172
Grade enrollment	17.2	28.3	18.8	22.5
Municipality inhabitants	11 784	11 684	12 305	12 585

Table A2: Robustness checks

Standardized values of national tests for fifth grade students 2007-2009					
SAMPLE	All students			Excluded student's with high probability of receiving special education.	
MODEL	Non-linearity model	Falsification test	Correction for student's sex	Schools with special education students	Schools with no special education students
Proportion of boys	-0.139			-0.0779	-0.211***
	(0.199)			(0.0494)	(0.0561)
Interaction					
Squared term	0.0688				
	(0.189)				
Falsification test.		0.0261			
		(0.0420)			
Peer measure corrected for individual student's sex			-0.104**		
			(0.0487)		
Constant	-0.867***				
	(0.0993)				
Observations	145,401	145,401	145,401	103,724	19,936
R-squared	0.097	0,097	0,097	0.096	0.109
Number of schools.	2,424	2,424	2,424		

Note: Controls include student gender, immigration status, parental education, parental income, family structure, number of brothers and sisters, birth order, grade enrollment, school type, proportion of male teachers, level of special education and number of assistants.. Robust standard errors. * significant at 10 % level. ** significant at 5 % level. *** significant at 1 % level.

Table A3: Balancing test. The relationship between gender composition and pre-determined student attributes.

	School fixed effects models					
Dependent variable	**Father's education**	**Mother's education**	**Dissolved families**	**Mother's earnings**	**Father's earnings**	**First generation immigrants**
Proportion of boys	-0.0140	0.0170	-0.00263	-5,297	10,805	-0.00253
	(0.0432)	(0.0442)	(0.0133)	(5,359)	(10,954)	(0.00274)
Constant	2.266***	2.621***	0.357***	65,648***	142,871***	0.0245***
	(0.0505)	(0.0490)	(0.0141)	(5,932)	(15,137)	(0.00319)
Observations	145,401	145,401	145,401	145,401	145,401	145,401
R-squared	0.189	0.243	0.050	0.120	0.048	0.016
Number of schools	2,424	2,424	2,424	2,424	2,424	2,424

Note: Controls include student gender, immigration status, parental education, parental income, family structure, number of brothers and sisters, birth order, grade enrollment, school type, proportion of male teachers, level of special education and number of assistants.. Robust standard errors. * significant at 10 % level. ** significant at 5 % level. *** significant at 1 % level.