Peer Effects in the Classroom: Evidence from a Natural Experiment

Bernardo Pincheira Sarmiento

bepincheirasarmiento@hse.ru

September 23, 2019

Abstract

The causal effect of classmates on students' academic performance, known as peer effects, is a question still open to debate in the field of education economics. In the paper I use the 8.8 earthquake that struck Chile in 2010 as a natural experiment for exogenous variation in peer composition, using the fact that the earthquake hit a random area of the country and forced some students to move into new schools for non-academic reasons. I use OLS and instrumental variables econometric specifications, with data of students observed in 2010, in affected or non-affected areas, to answer this question. My results show that the peer effects are positive and significant for students in both fourth and tenth grade. An increase of one standard deviation in the average score of the peers has an effect between 0.15 and 0.22 of a standard deviations in the student's own score. In addition, using IV quantile regressions, I find evidence of nonlinearities in the effect that potentially allow for within school sorting that might reduce inequality of performance between students and marginally increase average performance according to simulations.

JEL Classification: I21, I24, Q54.

Keywords: Peer effects, natural disasters, student achievement.

1 Introduction

The causal effect of classmates on a student's academic performance is a question still open to debate in the field of economics of education, and for policy makers. There is no consensus in the literature about the size of peer effects (for reviews see Sacerdote (2011), Epple and Romano (2011) and Sacerdote (2014)). The main contribution of the paper is: to identify the existence and size of peer effects in primary and upper secondary education in the class, taking into account selection into schools and possible non-random assignment into classes, to investigate the nonlinearities in these effects and to provide a policy recommendation on a possible way to improve efficiency within schools.

The problem of studying changes in peer quality is that the move of new students into a peer group is not always exogenous. A source of exogenous variation in peer composition is provided by natural disasters. One of the first studies to use an earthquake to obtain exogenous variation in the peer group is Cipollone and Rosolia (2007), for students in high school. Imberman et al. (2012) use hurricane Katrina as the source of variation and I use the earthquake that happened in Chile in 2010. Because of the earthquake, some students had to move to other schools either because their school was destroyed or because their parents had to move. Therefore, some classes received a share of displaced students for exogenous reasons. In my paper I focus on the effects on the stayers, students that did not change school after the earthquake. In this aspect, I follow a similar strategy to Angrist and Lang (2004), Imberman et al. (2012), Gibbons et al. (2013) among others, who focus on the peer effects on immobile students, at different group levels. Similar to Imberman et al. (2012), I use an instrumental variables strategy, with the proportion of displaced students in a class as an instrument for the average mean performance of peers (the average score of the students in the class discarding the score of the student). My first stage resembles Gibbons and Telhaj (2011), who look at the effect of incoming students on children who do not change school. The intuition is the following: given the magnitude and the timing of the earthquake, students had to be relocated into schools in a different way compared to children who move into another school because of academic reasons. The instrument is used to address what Manski (1993) called the "reflection" problem. To address selection into schools, which is one of the main econometric issues when estimating peer effects (Epple and Romano, 2011), I include school fixed effects, hence using within-school variation. Another potential source of bias is within school non-random allocation, when studying peer effects at the class level (Ammermueller and Pischke, 2009). One of the few published articles in the economics literature studying peer effects using data from Chile is McEwan (2003). Using a sample of students from eighth grade he shows that the results do not appear to be driven by within-school sorting. I follow his strategy to show that, based on parents' education, there is no evidence of a systematic allocation of students within schools. In addition to this information, as a further robustness check I remove from the sample schools in which there might be non-random allocation, based in the exam results. For other quasi-experimental settings see Sacerdote (2001) or Zimmerman (2003), who study peer effects in higher education.

My results show that peer effects are positive both in fourth and tenth grade. The statistical significance is somewhat sensitive to the specification used. In fourth grade, an increase of one standard deviation in the average score of the peers has an effect of approximately 0.16 standard deviations in the student's own score in mathematics and 0.15 standard deviations in Spanish in my baseline specification. For the median student it means moving approximately 5% up in the national distribution of scores. In tenth grade the effect in mathematics is such that the same improvement in the quality of peers would move the median student approximately 8% up in the distribution of scores.

There is some evidence for nonlinearities in the peer effect. In primary education, students near the median of the distribution of scores are on average less affected by their peers than other students who are either near the top or near the bottom of the class. On the other hand, the nonlinearities are less clear in the case of students in tenth grade, especially when analysing mathematics. One important result is that in all the groups analysed the peer effect is either positive or not statistically significantly different from zero.

The findings in the size of peer effects when using a linear-in-means model are in line with the current literature in primary education which study peer effects at the class level, in which the magnitudes range between modest (Burke and Sass, 2013) and no effects (Vigdor and Nechyba, 2007). In particular, my findings in Spanish are similar to the results in Ammermueller and Pischke (2009), who also relies on within-school variation, using a sample of European students in primary education. Furthermore, the slightly higher size of the effects and higher significance in tenth grade compared to fourth grade in mathematics is consistent to what Imberman et al. (2012) points out of the effects being larger for students in further stages in education. Moreover, results from Kang (2007) with students in seventh and eighth grade in mathematics show similar magnitudes to mine. In addition, coherent with the findings of Imberman et al. (2012) and my own, the results in Lavy and Schlosser (2011) show an increase in the magnitude of the effect in eighth grade compared to fifth grade. Contrary to the findings of Hanushek et al. (2003), I do find a statistically significant negative effect of the standard deviation of scores on achievement, although the magnitude is rather small. It is important to bear in mind that the variation induced by the earthquake, or another natural disaster, is different to the normal changes in peer composition (Imberman et al., 2012). I focus on the effect on the students who do not change school, most likely the least affected by the earthquake. Furthermore, the variation that I observe in peer composition even though is larger than in a normal year, is not as large as in Imberman et al. (2012), which may be positive when thinking about the external validity of my results.

The assessment used is a national standardised exam in mathematics and Spanish, taken every year around October by students in fourth grade (mostly aged between nine and ten years), including 2010 and every other year by students in tenth grade, also around October (fifteen to sixteen years), available for 2010 as well.

Related to my work is Tincani (2017) because the natural experiment used is the same, the earthquake that happened in Chile in 2010, although both the research question and the identification strategy differ. Tincani (2017) studies the effort choices in a class and shows that at least partly the peer effects seen on scores are due to rank concerns. In her study, the variation comes from the assumption that the intensity of the earthquake at a student's hometown increases their cost of effort to study. Not all the children in the same class live in the same town. Therefore, in some classes the dispersion in the cost of effort increases more than in other classes because students were exposed to the earthquake with different intensities. It is important to notice that in Tincani (2017) the classes studied are those which did not receive displaced students, so the variation comes from the different costs that children face for studying because of the earthquake. This is an important difference to my study, because the variation that I use comes from the difference in the share of displaced students in the class, closer to the strategy used by Imberman et al. (2012). Another important contrast is that while I focus exclusively on peer quality, Tincani (2017) also incorporates rank concerns in their model. On the other hand, while I study peer effects in 4th and 10th grade and report some differences between both groups, Tincani (2017) covers students in 8th grade only.

One study that looks for some sort of heterogeneity in peer effects is Patacchini et al. (2017). They distinguish the length of a friendship, either short or long, finding that only friendships that last more than one year have a persistent effect on academic outcomes. Another paper that investigates the heterogeneity in the peer effects is by Imberman et al. (2012), which looks for nonlinearities. They find monotonicity of peer effects and they reject the linear-in-means model.

A possible policy implication is that it might be feasible to reallocate students in a way that both improves average results and reduces inequality using the same resources, therefore improving efficiency.

However, my simulations show that the scope for improvement is rather limited for average performance, with an increase of 0.01 standard deviation in scores, and more promising for reducing inequality. However, this result should be taken with caution. For example, in the experiment made by Carrell et al. (2013) trying to group freshmen entering at the United States Air Force Academy in an optimal way, the result is that the treatment group performed worse than the group of peers allocated randomly. Similarly, Burke and Sass (2013) and Feld and Zolitz (2017) findings suggest that it may be detrimental for low-achieving students to have classmates of high ability. On the other hand, Carrell et al. (2013) and Feld and Zolitz (2017) studies focus on students in Higher Education, whereas in primary education it may be possible for the teacher to have a larger influence in the group formation via coursework for example.

The rest of the study is organised as follows: section 2 gives an overview of the Chilean school system and the Earthquake, section 3 covers the data, its sources and some descriptive statistics, section 4 is about the empirical strategy, in section 5 I go through the results while in section 6 I conclude.

2 The Chilean school system and the Earthquake

2.1 Chilean school system

The Chilean school system is organised in three tiers: primary school from first to fourth grade, lower secondary school from fifth to eighth grade, and upper secondary school from ninth to twelfth grade. Children usually start primary school at the age of six-seven years and leave upper secondary school at the age of seventeen-eighteen years.

There are three types of schools: public, private subsidised, and private schools (Valenzuela et al., 2014). The public, or municipal, are run by the municipalities, cannot charge any fees to the students and receive funding from the state. The private subsidised also receive funding from the state but in addition they were allowed to charge some tuition fees to the students at the time of the exams analysed in this study. Finally, private schools do not receive any funding from the state and are allowed to charge tuition fees. Around 90% of the students attend either public or subsidised schools, while less than 10% are enrolled in private schools in fourth grade. Public schools cannot select students before seventh grade. On the other hand, private subsidised and private schools may interview parents, hold some playing sessions for the children or do other type of selection processes (Anand et al., 2009).

In order for a student to progress to the next academic year, the following rules apply. From first to second grade and from third to fourth grade, every student with an attendance rate of at least 85% will be automatically promoted to the next year. The minimum attendance of 85% is a necessary but not sufficient condition in other grades. From second to third grade and from fourth to twelfth grade, students must either have passed all their subjects (minimum of 4.0 in a scale from 1 to 7 in teachers' assessment), have an average of 4.5 if they have failed one subject, or an average of 5.0 if they have failed 2 subjects. The averages are calculated considering the compulsory subjects in the national curricula, including Mathematics, Spanish, History among others. I refer to this average as GPA or teacher assessment in the rest of the paper.

To measure the quality of education, there are exams called "SIMCE", that are sat by students in fourth, eighth and tenth grade. In the case of students in fourth grade, this is every year, and for the other two cases only every other year.[1] The exams are centralised and the same for every child in the country. The results are publicly available at the school level. If a school shows a poor performance, then some sort of support is given to it to improve. On the other hand, if a school shows a good performance in the SIMCE, a bonus is paid to their teachers (Carnoy et al., 2007). According to Mizala and Urquiola (2009), the existing descriptive evidence does not suggest that the incentives have changed average testing performance.

In the case of primary school, the same teacher will teach most of the subjects in a given class. From fifth grade onwards, in most schools, students are taught each subject by a different teacher.

[1] In odd years students in eighth grade must sit the exam and in even years students in tenth grade must do the same.

2.2 Earthquake

Chile was struck by an earthquake on 27th February 2010, at 3:34 am local time. Its magnitude was 8.8 Richter in the epicentre, approximately 400 km south of Santiago, the capital, and it was the fifth-largest ever instrumentally recorded in the world (Astroza et al., 2010).[2] It affected six of the fifteen regions of the country. Approximately 80% of the population lived in that area. The earthquake occurred just 2 days before the official start of the school year, set to be the 1st of March 2010 and to include 38 weeks of teaching.[3] According to the information retrieved from the Chilean Ministry of Education (MINEDUC), more than half of all schools (4635 out of nearly 9000) were damaged. The damage was classified as: minor, moderate, or severe.

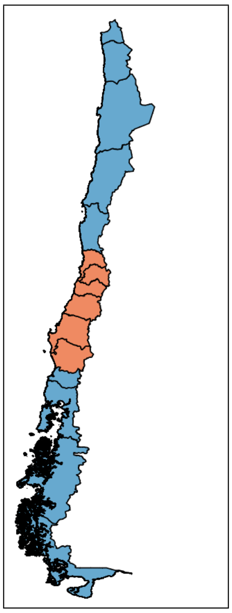
Figure 1 shows a map of the country divided by regions. The area affected by the earthquake includes 6 of the 15 regions, and is shown in red. The regions not affected by it are in blue.[4] It covers the three most populous cities in the country: Santiago, Valparaiso and Concepcion. The epicentre was on the coast, and after the earthquake there was a tsunami which was responsible of most of the deaths due to the earthquake. The official number

[2]Chile is a seismic country. According to the information from the National Seismological Centre (*Centro Sismologico Nacional*), before 2010 there were 109 earthquakes of magnitude at least 7.0, from Arica (northernmost part of the country) to Tierra del Fuego (southernmost part of the country). This, considering all the earthquakes from 1570. More accurate measures are from the beginning of the 20th century, and in that case the number of earthquakes before 2010 is 77. This represents on average 0.7 earthquakes per year, compared to 14.6 in the whole world since 1990 (or 0.07 per country per year, statistics from the United States Geological Survey). Another way to look at this is to say that Chile's surface is 0.56% of the world's, and its share of earthquakes is 5%.

[3]Information retrieved from MINEDUC.

[4]The definition of affected area comes from the Chilean Government, is at the regional level, and includes regions V-VI, VII, VIII, IX and the Metropolitan Region.

Figure 1: Map of Chile.



In red, the six regions affected by the earthquake. In blue, the areas not affected by it.

of deaths related to the earthquake and the subsequent tsunami was 525, according to the National Office for Emergencies ONEMI (*Oficina Nacional de Emergencias del Ministerio del Interior*).

3 Data

I use two sources of data. One is an administrative data from the Ministry of Education, which includes the following information at the student level: yearly academic performance measured by the teacher, a full record of schools attended every academic year, and town in which the school is located.[5] It covers from 2002 onwards. The second main source comes from the national standardised exam called SIMCE (System to Measure the Qual-

[5] A small proportion of students change school at the end of the academic year.

ity of Education), organised by the Agency for the Quality of Education. Its main purpose is to provide information about students' learning in different areas of the national curricula, and to contribute to the improvement of the quality and equity of education. The exam is sat every year by students in fourth grade (nine to ten years old), and every other year by students in eighth grade (thirteen to fourteen years old) and tenth grade (fifteen to sixteen years old). There are identifiers for the student, the classroom and the school. Therefore, it is possible to identify different classes within a school.

I use two datasets, one with information about students in fourth grade and the other one which has the same information about students in tenth grade, both in 2010, the year of the earthquake. I study those groups because they are highly comparable, they sat the exam at the same time and therefore one can expect that they were exposed in a similar way to the earthquake. Information about the students in both grades include the results in both mathematics and Spanish, gender, parents' education, household income, number of books available at home, class size, and a full record of the schools attended before.

From the information that comes from the dataset I create a dummy variable called *displaced*. The following criteria must be met for a student to be considered *displaced*: the school attended by the student in 2010 is different to the school attended in 2009, the year before the earthquake; the student was attending a school in 2009 in the affected area;[6] the student did not move

[6]There is information about the town in which the student resides but it is self-reported and in some cases inconsistent with the location of the school attended, therefore I use the location of the school instead.

to another school for academic reasons.[7] If all the previous conditions are met the variable *displaced* takes value 1, and value 0 otherwise. Therefore, children who were studying at schools outside the affected area in 2009 are not considered as displaced even if they did change school afterwards. Data about the school attended before and after the earthquake are in Table 1 for students in fourth grade and Table 2 for students in tenth grade. In both grades approximately 60% of the changes are between schools of the same type. From the rest, the majority of the changes are between Municipal and Private Subsidised schools.

Table 1: Displaced students 4th grade

Type of school before and after the earthquake	**Frequency**	**%**
Municipal to Municipal	2,829	24.17
Municipal to Subsidised	1,950	16.66
Municipal to Private	16	0.14
Subsidised to Subsidised	3,929	33.56
Subsidised to Municipal	2,091	17.86
Subsidised to Private	230	1.96
Private to Private	368	3.14
Private to Municipal	28	0.24
Private to Subsidised	265	2.26
Total	**11,706**	**100**

Summary statistics of the main variables used in the regression are shown in Table 3 and Table 4. The first two variables are the raw score of the students in mathematics and the class standard deviation in mathematics. The next two variables are the same but for Spanish. GPA is the raw score for teacher assessment, which can go from 1 to 7. Father education and Mother

[7]There is a survey for parents in which they are asked about the main reason of choosing the current school at which the student is.

Table 2: Displaced students 10th grade

Type of school before and after the earthquake	**Frequency**	**%**
Municipal to Municipal	428	10.67
Municipal to Subsidised	523	13.04
Municipal to Private	23	0.57
Subsidised to Subsidised	1,720	42.87
Subsidised to Municipal	588	14.66
Subsidised to Private	139	3.46
Private to Private	389	9.70
Private to Municipal	26	0.65
Private to Subsidised	176	4.39
Total	**4,012**	**100**

education indicate how many years of education received their parents, and Female is a dummy which takes value 1 if the student is female and 0 if the student is male. Finally, I include a dummy variable which indicates if there are more than 10 books at the student's home, and I also include a measure of household monthly income, which includes 13 categories, each category separated by approximately £150. In my regressions I use standardised scores for both the SIMCE exam and teacher's assessment, with mean 0 and standard deviation of 1.

The separation between a change of school caused by academic vs non-academic reasons is relevant. In the cases where parents say that the reason of the change was because they found a better school, average score in the SIMCE exam in mathematics of the new school is on average 20 points higher, nearly a third of one standard deviation, compared to the previous school attended by the student. On the other hand, for the subgroup of students who are considered displaced, the current school's score is on average 2 points lower (approximately 0.04 standard deviations) than in the previous school.

Table 3: Summary statistics for students in fourth grade, 2010

Variable	**Mean**	**Std. Dev.**
Maths raw score	262.932	52.374
Class SD in maths scores	43.784	6.908
Spanish raw score	278.674	48.364
Class SD in Spanish scores	43.083	8.000
Female	0.511	0.5
GPA 1st grade	6.357	0.465
GPA 2nd grade	6.187	0.492
GPA 3rd grade	6.005	0.517
Class size	33.054	7.285
More than 10 books at home	0.687	0.464
Father education	12.01	3.439
Mother education	12.108	3.555
Household monthly income	4.714	3.404
Observations	107095	

3.1 Subsample analysed

From the total of more than 200,000 students observed in each exam, approximately 50% are in the main analysis. First, in order to include controls as parents' education, I restrict the sample to observations in which the parents' survey was answered, which reduces the sample size by approximately 10%. Also, I only considered students who didn't repeat any grade, reducing my sample. However, in the peer group I do consider all the students who have a valid score. Hence, if there is a classroom with 30 students, including one who repeated some level before and another one for whom I do not have additional information, the final sample will consist of 28 students, but for the peer group I will consider all the 30 students. That is, when calculating

Table 4: Summary statistics for students in tenth grade, 2010

Variable	**Mean**	**Std. Dev.**
Maths raw score	274.794	63.247
Class SD in maths scores	42.754	7.681
Spanish raw score	273.167	51.251
Class SD in Spanish scores	38.915	6.015
Female	0.529	0.499
GPA 7th grade	5.767	0.551
GPA 8th grade	5.788	0.545
GPA 9th grade	5.534	0.548
Class size	33.978	6.964
More than 10 books at home	0.803	0.398
Father education	11.897	3.697
Mother education	11.993	3.79
Household monthly income	5.188	3.56
Observations	89788	

the average peer score, I include all the students who have a valid score in the exam, even if they are not included in the final sample. In addition, there are some students, especially in rural areas, who have less than 10 classmates, so I restricted my sample to students in a class with at least 10 students.[8] Furthermore, to use school fixed effects I had to restrict the sample to students in schools with at least two classes per grade, otherwise there would be no variation left within schools with only one class. Therefore, my final sample I observe 107,095 students in fourth grade and 89,788 students in tenth grade.

4 Empirical Strategy

The aim of the paper is to identify the causal effect of peers on students' academic achievement, measured with national standardised exam scores in

[8] I tried different thresholds, also having a minimum of 15 or 20 students, without any significant variation in the main results.

mathematics an Spanish. I start with a standard linear-in-means OLS specification for peer effects in the class:

$$Y_{ics} = \beta_0 + \beta_1 P_{-ics} + \beta_2 X_{ics} + \beta_3 M_{cs} + k_s + \epsilon_{ics} \tag{1}$$

where Y_{ics} is the Spanish or mathematics score in SIMCE for student i in class c, attending school s, P_{-ics} is the average peer performance of all the students excluding student i in class c, attending school s, X_{ics} are observable characteristics of the student, M_{cs} are observable characteristics of the class, k_s are school fixed effects and ϵ_{ics} is the error term that captures unobserved factors that determine academic achievement.

This specification follows what McEwan (2003) does with Chilean data as I do but with a different cohort and grade. He argues that including school-specific fixed effects can help to at least partially address the bias in the estimates of β_1 caused by families with similar unobserved preferences choosing similar schools, commonly referred as selection into schools (see Manski (1993) and Sacerdote (2014)). One may think that students are not randomly allocated into schools. It may be more likely that if parents can choose, they will try to put their children in a "good" school instead of a "bad" one, leading to self-selection of students into peer groups.[9] Therefore, it is important to address sorting into schools. This selection problem is particularly important in the Chilean context, which has a high socioeconomic segregation (Valenzuela et al., 2014). A possible improvement with

[9] What is good or bad will vary from case to case but one possible measure could be the performance of students in that school in a national exam.

respect to McEwan (2003) is that I include information of previous ability for each student, measured by teacher's assessment. Furthermore, equation 1 is similar to the linear-in-means specification used in Imberman et al. (2012). The main difference is that I use the class as the relevant peer group and their study uses the grade-school as the relevant peer group (for other examples using instruments at the grade-school level see Lavy and Schlosser (2011) and Lavy et al. (2011)). In the Chilean education system, most of the interactions, especially in mathematics and Spanish, occur at the class level rather than at the school level. Moreover, in primary education the head teacher usually covers both subjects for a given class. Hence, I opt to group at the class level as in McEwan (2003) and Tincani (2017), both using Chilean data. Nevertheless, even when including school fixed effects and a measure of previous ability the estimates of β_1 may be biased because of the "reflection" problem (Manski, 1993). The reflection problem arises because the score of student i has an impact on the average score of her peers, which may cause a bias in the β_1 coefficient. Therefore, the next step is to find an instrument that generates exogenous variation into the peer composition.

To add this exogenous variation, the main specification to look for peer effects uses the fraction of displaced students in the class as an instrument for average peer performance, in a similar way as Imberman et al. (2012). The identification strategy relies on the following assumptions: the random allocation of students into schools and within schools because of the earthquake.[10] Under this assumption, the fraction of displaced students is an exogenous

[10] For the students considered as displaced, the quality of the school attended in 2009 is on average similar to the quality of the school attended in 2010. This is not the case for students who move for academic reasons, which I exclude from the sample.

source of variation for the peer composition, addressing the reflection problem when identifying peer effects. Therefore, the first stage regression can be written as:

$$P_{-ics} = \alpha_0 + \alpha_1 \frac{D_{cs}}{N_{cs}} + \alpha_2 X_{ics} + \alpha_3 M_{cs} + k_s + e_{ics} \tag{2}$$

where D_{cs} is the number of displaced students and N_{cs} is the number of students in class c, attending school s. Consequently, the fraction of displaced students will not be necessarily be the same for two different classes within a school.

The second stage regression is:

$$Y_{ics} = \beta_0 + \beta_1 \hat{P}_{-ics} + \beta_2 X_{ics} + \beta_3 M_{cs} + k_s + \epsilon_{ics} \tag{3}$$

where $\hat{P}_{-ics}$ is the predicted average peer score for all the students excluding student i in class c, attending school s.

4.1 Selection into classes

The identification strategy relies on the assumption that within schools, children are allocated randomly into different classes. If it is not the case, then my results may be biased. A preliminary check that I perform is to compare the average peer score when using either the class or the school as the aggregation group. For all the grades and subjects the correlation is above 0.9, suggesting no large differences within a school. A more formal test is done by McEwan (2003). In his study he compares a measure of average mothers' education at the classroom level with a simulated random class allocation. His findings suggest that the differences are not large enough to suggest a systematic assignment into classes by ability. I follow his approach and I reach the same conclusions for both cohorts studied. My results are in Table 5, and includes the same procedure using fathers' education as well. The observed and the simulated dispersion are virtually the same across grades and subjects. As an additional way to attempt to rule out the problem of non-random allocation within schools, I compare outcomes of different classes within a school. The method I used to test this claim is by looking at a subsample of students which are in schools that do not appear to set their classes by ability. The criteria that I impose is that the difference in scores between classes must be not statistically significant either in mathematics or Spanish. If a school does not meet it, then I remove it from the subsample. Therefore, I keep in my sample only schools in which their classes have a similar score in both mathematics and Spanish. Finally, I perform the regressions for the main sample, including all the students as described in the data section, and a subsample of students in schools that meet the previous criteria of no differences that are statistically significant. The latter approach has some limitations. First, it uses the current outcome in mathematics or Spanish as an indicator of an allocation into different groups that happened in the past. Second, by construction it reduces the within-school variation, which is the source of my identification strategy. [11]

[11] This subsample is approximately 80% of the students used in the main regressions.

Table 5: Selection into classes according to family characteristics

Variable	Observed Std. Dev.	Simulated Std. Dev.
Class Father Ed. 10th grade	2.194	2.193
Class Mother Ed. 10th grade	2.025	2.025
Class Father Ed. 4th grade	2.043	2.179
Class Mother Ed. 4th grade	1.930	2.065

4.2 Nonlinearities

There is a growing literature looking beyond linear-in-means peer effects (Tincani, 2017). In this study I use IV quantile regressions to show how the peer effect vary depending on the position of student i in the distribution of scores. The focus of this section is to estimate an equivalent version of equation 3, the specification with instrumental variables, for 10 different quantiles, using the estimator developed by Powell (2017). The importance of nonlinearities in peer effects lies in the fact that without them, it is not possible to find any allocation of students that increases average performance (see Carrell et al. (2013) for a previous attempt) based on peer quality, because the total amount of ability is fixed. On the other hand, other targets such as reducing educational inequality or increasing the performance of certain students are possible to achieve even in the absence of nonlinearities in the importance of peers.

5 Results

5.1 Main results

The main results of the effect of peers on students' performance are summarised in four tables, one for each combination of cohort-outcome. The combinations are: mathematics in fourth grade (M4), Spanish in fourth grade (S4), mathematics in tenth grade (M10) and Spanish in tenth grade (S10). The sample analysed comprises students who did not move to another school between 2009 and 2010, attending schools with at least two different classes and at least ten students per classroom. In columns 1-3 are the results of estimating equation 1 and column 4 shows the results when estimating equation 3.

The first column of Table 6 for M4 shows that the coefficient of peer effects is positive and significant when including only individual controls. The magnitude is 0.420 and it is significant at the 1% level. In line with the literature, the coefficient of -0.223 for the dummy variable Female shows that male students perform better than their female counterparts in mathematics. In column two, when including class level controls and school fixed effects, the coefficient of peer effects nearly halves to 0.220 and remains significant at the 1% level, showing that selection into schools is an important driver of upward biases in the results when not taken into account. In column 3, when adding controls from the family background of the student, the results are virtually unchanged. Column 4 has the same variables as column three but shows the results when estimating equation 3, the instrumental variables specification. The main difference compared to column three is that the coefficient of peer effect decreases to 0.150 and is significant only at the 5% level.

Table 7 shows the results for the group M10. In line with the results in mathematics for students in primary education, the coefficient of peer effects decreases when including school fixed effects, from 0.584 to 0.266, always significant at the 1% level. Furthermore, when using instrumental variables, the coefficient for peer effects further declines to 0.219, with the same statistical significance as before. On the other hand, the gender gap marginally increases.

The results of Spanish for students in fourth grade are in Table 8. One substantial difference that the coefficient for the variable Female, which now is positive and significant at the 1% level. The coefficient for peer effects in column 1 is 0.326, significant at the 1% level. It goes down to 0.126 in column two when controlling for selection into schools but remains significant at the 1% up to column 3. In column four, when estimated using instrumental variables it increases to 0.141, but due to the larger standard error, it is not statistically significantly different from the coefficient in column three estimated with OLS. Overall, results in Table 8 tell a similar story as those for mathematics in fourth grade.

Finally, the results of Spanish in tenth grade, which I include in Table 9, do not change substantially when compared with Spanish in fourth grade. The peer effect coefficient is 0.487 and significant at the 1% level when using OLS and individual controls, decreases to 0.181 when accounting for selection into schools and further decreases to 0.170 after adding some family background characteristics. In column 4, when using instrumental variables, the coefficient for peer effects decreases to 0.124, and it is significant only at the 10% level. One noticeable difference when comparing the results with those for students in fourth grade in Spanish is that the gender gap reverses, with male students now performing slightly better than their female counterparts, after controlling for observables.

The results are consistent with the problems for estimating peer effects mentioned in Manski (1993). I show that the main issue is selection into schools, and the reduction in the coefficient of the peer effect when estimated with instrumental variables may arise because it deals with the reflection problem. My results are consistent with the findings of Imberman et al. (2012), who show lower coefficients when using IV compared to OLS, even though in my case they remain significant, at least at the 10%. This in turn may be explained by the fact that I look for peer effects at the class level, as in Gibbons et al. (2013), and it is plausible that the influence of classmates may be higher than the influence of other students in the same grade but not in the same class.

The goodness of fit, reported at the end of each table for the results estimated with OLS, increases when including information about schools and class level controls but remains virtually the same when adding the family controls in column 3.

5.2 Robustness checks

One important issue when trying to get causal estimates of peer effects is the selection into the peer group under study. Even though I argue that selection

Table 6: Peer effects on stayers, OLS and IV (4th grade, Mathematics)

	(1) OLS	(2) OLS	(3) OLS	(4) IV
Peer score Maths	0.420***	0.220***	0.217***	0.150**
	(0.004)	(0.009)	(0.009)	(0.074)
Female	-0.223***	-0.267***	-0.264***	-0.263***
	(0.005)	(0.005)	(0.005)	(0.005)
Individual controls	Yes	Yes	Yes	Yes
Class level controls	No	Yes	Yes	Yes
School FE	No	Yes	Yes	Yes
Family background controls	No	No	Yes	Yes
Observations	107095	107095	107095	107095
R^2	0.519	0.580	0.582	N/A

Standard errors in parentheses clustered at the class level. Column 1 includes individual measures of previous academic performance measured by the teacher. Column 2 adds class size, class size squared and school fixed effects. Column 3 adds characteristics from the family background including parents' education, books available at home and household monthly income. Column 4 includes the same controls as column 3 but shows the results of the IV estimation, using the shared of displaced students in the class as an instrument for mean ability. F statistic of weak identification for IV: 49

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Peer effects on stayers, OLS and IV (10th grade, Mathematics)

	(1) OLS	(2) OLS	(3) OLS	(4) IV
Peer score Maths	0.584***	0.266***	0.258***	0.219***
	(0.003)	(0.009)	(0.009)	(0.050)
Female	-0.249***	-0.278***	-0.275***	-0.275***
	(0.005)	(0.005)	(0.005)	(0.005)
Individual controls	Yes	Yes	Yes	Yes
Class level controls	No	Yes	Yes	Yes
School FE	No	Yes	Yes	Yes
Family background controls	No	No	Yes	Yes
Observations	89788	89788	89788	89788
R^2	0.638	0.672	0.674	N/A

Standard errors in parentheses clustered at the class level. Column 1 includes individual measures of previous academic performance measured by the teacher. Column 2 adds class size, class size squared and school fixed effects. Column 3 adds characteristics from the family background including parents' education, books available at home and household monthly income. Column 4 includes the same controls as column 3 but shows the results of the IV estimation, using the shared of displaced students in the class as an instrument for mean ability. F statistic of weak identification for IV: 66

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Peer effects on stayers, OLS and IV (4th grade, Spanish)

	(1) OLS	(2) OLS	(3) OLS	(4) IV
Peer score Spanish	0.326***	0.126***	0.124***	0.141**
	(0.004)	(0.008)	(0.008)	(0.061)
Female	0.088***	0.070***	0.073***	0.073***
	(0.005)	(0.005)	(0.005)	(0.005)
Individual controls	Yes	Yes	Yes	Yes
Class level controls	No	Yes	Yes	Yes
School FE	No	Yes	Yes	Yes
Family background controls	No	No	Yes	Yes
Observations	107095	107095	107095	107095
R^2	0.386	0.438	0.440	N/A

Standard errors in parentheses clustered at the class level. Column 1 includes individual measures of previous academic performance measured by the teacher. Column 2 adds class size, class size squared and school fixed effects. Column 3 adds characteristics from the family background including parents' education, books available at home and household monthly income. Column 4 includes the same controls as column 3 but shows the results of the IV estimation, using the shared of displaced students in the class as an instrument for mean ability. F statistic of weak identification for IV: 51

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Peer effects on stayers, OLS and IV (10th grade, Spanish)

	(1) OLS	(2) OLS	(3) OLS	(4) IV
Peer score Spanish	0.487***	0.181***	0.170***	0.124*
	(0.003)	(0.008)	(0.008)	(0.065)
Female	-0.036***	-0.026***	-0.022***	-0.022***
	(0.005)	(0.005)	(0.005)	(0.005)
Individual controls	Yes	Yes	Yes	Yes
Class level controls	No	Yes	Yes	Yes
School FE	No	Yes	Yes	Yes
Family background controls	No	No	Yes	Yes
Observations	89788	89788	89788	89788
R^2	0.511	0.544	0.548	N/A

Standard errors in parentheses clustered at the class level. Column 1 includes individual measures of previous academic performance measured by the teacher. Column 2 adds class size, class size squared and school fixed effects. Column 3 adds characteristics from the family background including parents' education, books available at home and household monthly income. Column 4 includes the same controls as column 3 but shows the results of the IV estimation, using the shared of displaced students in the class as an instrument for mean ability. F statistic of weak identification for IV: 43

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

does not seem likely to be the driver of my results, I attempt further checks. To start, I perform the same regressions for a subsample of students who are in schools where the different classes within a grade do not show differences that are statistically significantly differently from zero neither in mathematics nor in Spanish. The results change for students in fourth grade, especially when using instrumental variables. In the case of M4 reported in Table 10, the coefficient of peer effects remains positive and significant at the 1% level in columns 1-3, with a similar order of magnitude to the baseline results in Table 6. However, when using instrumental variables, the coefficient of peer effect is no longer statistically significantly different from zero. For the case of S4 reported in Table 12, columns 2 and 3 remain positive and significant at the 1% level, although with a lower magnitude compared to the results in the main sample. Similarly to the case of mathematics, when using instrumental variables the coefficient is not significant any more. It is important to take these results with caution, because these results may arise from the loss in variation between classes within a school rather than showing a systematic selection of students into classes in primary school, under the current rules.

The differences are smaller when looking at tenth grade, group M10 in Table 11 and group S10 in Table 13. The magnitudes and significance levels are virtually the same as those presented in the main sample. One potential explanation to the difference between grades is that in primary education each class is usually taught most of the subjects by their head teacher, including mathematics and Spanish. This potentially means a larger variation of scores within a school, showing a spurious larger selection of students within schools. It is documented in the literature that there are differences in teacher effectiveness, even when controlling for certification status and experience as in Kane et al. (2008). Moreover, in some cases teacher credentials are not positively reflected in student's academic achievement (see Carrell and West (2010)). In a recent review Coenen et al. (2018) find that teacher experience does matter for students' test scores but not general teacher certifications. On the contrary, in secondary education usually every class in the same grade within a school will have the same teacher for mathematics and another one for Spanish. This fact may explain the lower variation within schools for secondary schools.

Overall the results suggest that the issue of selection within schools might exist, but that it is not the main driver for the results in peer effects, which is also what McEwan (2003) finds when using a cohort of eighth graders.

Another concern is that mean ability may not be the only aspect that matters when quantifying peer effects. It is plausible that dispersion in ability could be relevant as well. In fact, the evidence about the effects of heterogeneous class composition is mixed (Cortes et al., 2014). Therefore, I add the standard deviation of class scores in my regressions, as a further robustness check. The results are in Tables 14-17 in the appendix section. The standard deviation of scores within a class has a negative and statistically significant association with students' performance in both grades and both subjects. Increasing the dispersion in one standard deviation has an effect of reducing students' scores by less than 0.02 standard deviations, everything else equal. Besides this effect, it does not affect largely the results in mathematics, with the peer effects remaining with similar magnitudes and statistical significance compared to the baseline results. However, it does change some results for

Table 10: Peer effects on stayers, no large differences between classes (4th grade, Mathematics)

	(1) OLS	(2) OLS	(3) OLS	(4) IV
Peer score Maths	0.416***	0.141***	0.139***	-0.187
	(0.004)	(0.013)	(0.013)	(0.220)
Female	-0.223***	-0.267***	-0.264***	-0.262***
	(0.005)	(0.005)	(0.005)	(0.005)
Individual controls	Yes	Yes	Yes	Yes
Class level controls	No	Yes	Yes	Yes
School FE	No	Yes	Yes	Yes
Family background controls	No	No	Yes	Yes
Observations	88693	88693	88693	88693
R^2	0.513	0.579	0.581	N/A

Standard errors in parentheses clustered at the class level. Column 1 includes individual measures of previous academic performance measured by the teacher. Column 2 adds class size, class size squared and school fixed effects. Column 3 adds characteristics from the family background including parents' education, books available at home and household monthly income. Column 4 includes the same controls as column 3 but shows the results of the IV estimation, using the shared of displaced students in the class as an instrument for mean ability. F statistic of weak identification for IV: 13

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Peer effects on stayers, no large differences between classes (10th grade, Mathematics)

	(1) OLS	(2) OLS	(3) OLS	(4) IV
Peer score Maths	0.584***	0.263***	0.255***	0.221***
	(0.003)	(0.010)	(0.009)	(0.056)
Female	-0.249***	-0.279***	-0.276***	-0.277***
	(0.005)	(0.005)	(0.005)	(0.005)
Individual controls	Yes	Yes	Yes	Yes
Class level controls	No	Yes	Yes	Yes
School FE	No	Yes	Yes	Yes
Family background controls	No	No	Yes	Yes
Observations	83577	83577	83577	83577
R^2	0.641	0.675	0.676	N/A

Standard errors in parentheses clustered at the class level. Column 1 includes individual measures of previous academic performance measured by the teacher. Column 2 adds class size, class size squared and school fixed effects. Column 3 adds characteristics from the family background including parents' education, books available at home and household monthly income. Column 4 includes the same controls as column 3 but shows the results of the IV estimation, using the shared of displaced students in the class as an instrument for mean ability. F statistic of weak identification for IV: 47

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 12: Peer effects on stayers, no large differences between classes (4th grade, Spanish)

	(1) OLS	(2) OLS	(3) OLS	(4) IV
Peer score Spanish	0.320***	0.040***	0.039***	-0.069
	(0.004)	(0.012)	(0.012)	(0.155)
Female	0.085***	0.067***	0.069***	0.070***
	(0.006)	(0.006)	(0.006)	(0.006)
Individual controls	Yes	Yes	Yes	Yes
Class level controls	No	Yes	Yes	Yes
School FE	No	Yes	Yes	Yes
Family background controls	No	No	Yes	Yes
Observations	88693	88693	88693	88693
R^2	0.378	0.435	0.436	N/A

Standard errors in parentheses clustered at the class level. Column 1 includes individual measures of previous academic performance measured by the teacher. Column 2 adds class size, class size squared and school fixed effects. Column 3 adds characteristics from the family background including parents' education, books available at home and household monthly income. Column 4 includes the same controls as column 3 but shows the results of the IV estimation, using the shared of displaced students in the class as an instrument for mean ability. F statistic of weak identification for IV: 14

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13: Peer effects on stayers, no large differences between classes (10th grade, Spanish)

	(1) OLS	(2) OLS	(3) OLS	(4) IV
Peer score Spanish	0.485***	0.165***	0.155***	0.131*
	(0.003)	(0.009)	(0.009)	(0.073)
Female	-0.037***	-0.026***	-0.022***	-0.022***
	(0.005)	(0.006)	(0.006)	(0.006)
Individual controls	Yes	Yes	Yes	Yes
Class level controls	No	Yes	Yes	Yes
School FE	No	Yes	Yes	Yes
Family background controls	No	No	Yes	Yes
Observations	83577	83577	83577	83577
R^2	0.513	0.547	0.551	N/A

Standard errors in parentheses clustered at the class level. Column 1 includes individual measures of previous academic performance measured by the teacher. Column 2 adds class size, class size squared and school fixed effects. Column 3 adds characteristics from the family background including parents' education, books available at home and household monthly income. Column 4 includes the same controls as column 3 but shows the results of the IV estimation, using the shared of displaced students in the class as an instrument for mean ability. F statistic of weak identification for IV: 32

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

students in Spanish for students in tenth grade, as I show in Table 17. The peer effect is no longer significant at the 10% level (the p-value, not reported in the table, is approximately 0.11).

Furthermore, given the characteristics of the earthquake, students who received a large influx of displaced new classmates may have been affected not only by the changes in peer composition but also by a radical change in the number of students in the class. I test my results by adding growth in class size (as percentage) to the controls, in Tables 18-21. The coefficients of peer effects estimated with instrumental variables are lower in magnitude and, with the exception of mathematics in tenth grade, are not statistically significantly different from zero anymore. On the other hand, the coefficients of peer effects remain stable when I use OLS estimation. The results show that an increase in class size of 10% reduces the score of a student between 0 and 0.02 standard deviations.

5.3 Nonlinearities

My results using quantile regressions suggest some non-linearity in the peer effects, especially in mathematics with different patterns when comparing fourth grade with tenth grade. I show these features graphically in Figure 2. For the group M4, it shows that students in the middle of the distribution of scores are less affected by their peers than students towards either the bottom or the top of the distribution of scores. The differences are statistically significant. For the group M10 the nonlinearities are of a smaller magnitude, except for students in the bottom of the distribution of scores being clearly less affected by their peers than the rest. For Spanish, the opposite is true for the group S4, in which students at the bottom of the distribution seem more affected by their peers than the rest. Finally, for the group S10 the most interesting pattern is that one quantile towards the top achievers is affected negatively by their peers, even though the coefficient is not statistically significantly different from zero.

One interesting implication from the results of the group M4 is that it may be possible to reallocate the students within a school in a way that marginally improves the average score of the school and reduces the inequality between students' outcomes. The alternative allocation is to put the bottom 20% of students with the top 30% in one group, and have a second group with the rest. To test this claim I run some simulations. The typical school in my sample has 2 classes and on average 30 students in each class, grouped in a way that suggests a random allocation within the school. I run the following simulation: first I generate randomly 60 observations drawing scores from a normal distribution with mean 250 and standard deviation of 50, which are approximately the parameters of the SIMCE exam. Then I standardise their scores to mean zero and standard deviation of 1. I calculate the peer score for student i considering all the students in the school except the student's own score. This would be an equivalent to a perfectly mixed peer group, similar to what I observe in the data. Then I rank the observations according to the scores and I put in one group the lowest 12 observations with the top 18 and in the other group the other 30 observations. I recalculate the peer score for each observation and, using the results from the quantile regressions, I recalculate the score for each observation with the simulated new peer score. On average this new allocation improves the score of the group by approximately 0.01 standard deviations. The magnitude is similar to an increase of one year in father's education for example.

One study that tries to allocate groups in an optimal way is Carrell et al. (2013). They group freshmen entering at the United States Air Force Academy in an way supposed to benefit those who performed worst in exams. However, the result is that the treatment group performed worse than the group of peers allocated randomly. The explanation is that freshmen formed subgroups of more homogeneous peers, for example among students who performed badly in previous exams. One difference between their experiment and the context studied in this paper is that in primary education it is possible for the teacher to allocate students in sub-groups either for working in the classroom or at home. If done in an earlier stage of the academic year, it may be possible form heterogeneous groups of students which are beneficial for the low achievers. I am aware of the difficulties of identifying in an early stage the ability of children. Another difference is that according to the information they had before the experiment, it should have helped the low achievers to improve without negatively affecting the high achievers. My proposal is less ambitious in that sense because it may decrease the score of the high achievers. However, this effect is more than offset in the simulations by the improvement of the low achievers, improving average performance and reducing inequality.

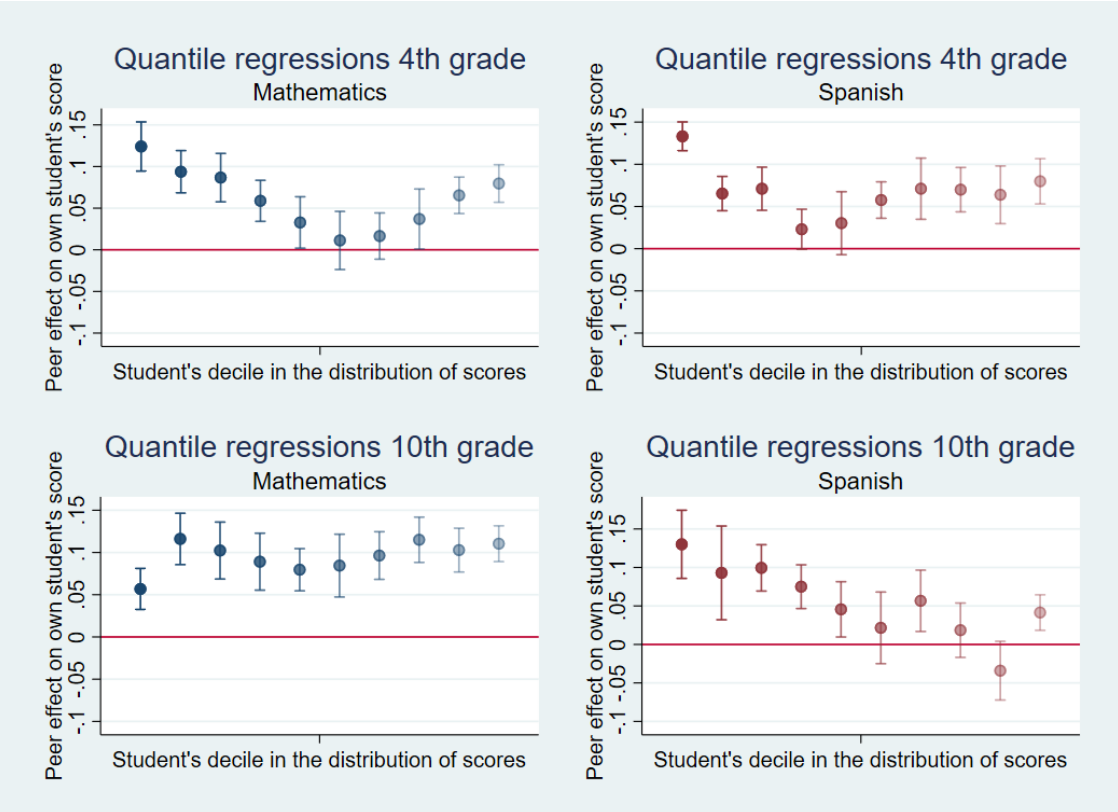


Figure 2: Quantile regressions using instrumental variables and school fixed effects.

6 Conclusion

The first conclusion is that apparently the largest source of bias of the peer effects comes from the self-selection into schools. Across all the four specifications the coefficient of peer effect decreases by nearly 50% or more once controlling for school fixed effects and class characteristics There is a further reduction in the coefficient when using instrumental variables, but the difference is, on average, smaller in magnitude. When taking into account a possible sorting within schools, some coefficients go further down but not as much as when controlling for school fixed effects. Secondly, the main difference between students in fourth grade compared to those in tenth grade is that the peer effect in mathematics is consistently higher for older students. The same is not true for Spanish, in which case it depends in the specification. Another relevant conclusion is that the results derived from the estimation using instrumental variables are more sensitive to changes in the sample analysed or the controls included.

Regarding nonlinearities, the results suggest that peer effects are not linear. For both subjects in fourth grade, students in the middle of the distribution of scores are less affected by their peers than students in the bottom of the same distribution. Furthermore, only one of the quantile groups have a negative point estimate for the peer effect, even though this is not statistically significantly different from zero. I can conclude that on average it is better to have better peers, or at least it is not harmful. However, the nonlinearities found are not strong enough to allow me to find a pareto improving allocation of students within schools. I do show some results derived from simulations that give some hope to reduce educational inequality without negatively affecting average achievement, and even improving it marginally.

An important caveat of this study is that displaced students may have affected the stayers through channels not controlled by my empirical strategy. For example, either student behaviour or class size might have been affected in a different way across classes within a school, which may in turn introduce some sort of bias to the main results. Another possible limitation of the study may come from the issue of selection into classes, which may not be entirely ruled out. Another limitation may be the external validity of my results, given that I use a natural disaster as a source of variation.

Patacchini et al. (2017) finds that only long-lasting friendships have an important effect. Even though the data does not allow me to identify groups of friends within the same class, it may be an interesting extension of this study to consider the length of the student in the same peer group.

References

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

P. Anand, A. Mizala, and A. Repetto. Using school scholarships to estimate the effect of private education on the academic achievement of low-income students in Chile. *Economics of Education Review*, 28:370–381, 2009.

J.D. Angrist and K. Lang. Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. *The American Economic Review*, 94(5):1613–1634, 2004.

M. Astroza, S. Ruiz, and R. Astroza. Damage Assessment and Seismic Intensity Analysis of the 2010 (Mw 8.8) Maule Earthquake. *Earthquake Spectra*, 28(S1):S145–S164, 2010.

M. Burke and T. Sass. Classroom Peer Effects and Student Achievement. *Journal of Labor Economics*, 31(1):51–82, 2013.

M. Carnoy, I. Brodziak, A. Molina, and M. Socias. The limitations of teacher pay incentive programs based on inter-cohort comparisons: The case of Chile's SNED. *Education Finance and Policy*, 2(3):189–227, 2007.

S. Carrell and J. West. Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors. *Journal of Political Economy*, 118(3):409–432, 2010.

S. Carrell, B. Sacerdote, and J. West. From natural variation to optimal policy? The importance of endogenous peer group formation. *Econometrica*, 81(3):855–882, 2013.

P. Cipollone and A. Rosolia. Social Interactions in High School: Lessons from an Earthquake. *The American Economic Review*, 97(3):948–965, 2007.

J. Coenen, I. Cornelisz, W. Groot, H. Maassen van den Brink, and C. Klaveren. Teacher Characteristics and their Effects on Student Test Scores: A Systematic Review. *Journal of Economic Surveys*, 32(3):848–877, 2018.

K. Cortes, , and J. Goodman. Ability-Tracking, Instructional Time, and Better Pedagogy: The Effect of Double-Dose Algebra on Student Achievement. *American Economic Review*, 104(5):400–405, 2014.

D. Epple and R. Romano. Peer Effects in Education: A Survey of the Theory and Evidence. *Handbook of Social Economics*, 1B:1053–1163, 2011.

J. Feld and U. Zolitz. Understanding Peer Effects: On the Nature, Estimation and Channels of Peer Effects. *Journal of Labor Economics*, 35(2):387–428, 2017.

S. Gibbons and S. Telhaj. Pupil mobility and school disruption. *Journal of Public Economics*, 95:1156–1167, 2011.

S. Gibbons, O. Silva, and F. Weinhardt. Everybody needs good neighbours? Evidence from students outcomes in England. *The Economic Journal*, 123: 831–874, 2013.

E. Hanushek, J. Kain, J. Markman, and S. Rivkin. Does peer ability affect student achievement? *Journal of Applied Econometrics*, 18:527–544, 2003.

S. Imberman, A. Kugler, and B. Sacerdote. Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees. *The American Economic Review*, 102(5):2048–2082, 2012.

T. Kane, J. Rockoff, and D. Staiger. What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review*, 27:615–631, 2008.

C. Kang. Classroom peer effects and academic achievement: Quasi-randomization evidence from South Korea. *Journal of Urban Economics*, 61:458–495, 2007.

V. Lavy and A. Schlosser. Mechanisms and Impacts of Gender Peer Effects at School. *American Economic Journal: Applied Economics*, 3:1–33, 2011.

V. Lavy, D. Paserman, and A. Schlosser. Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom. *The Economic Journal*, 122:208–237, 2011.

C. Manski. Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, 60(3):531–542, 1993.

P. McEwan. Peer effects on student achievement: evidence from Chile. *Economics of Education Review*, 22:131–141, 2003.

A. Mizala and M. Urquiola. School markets: The impact of information approximating schools' effectiveness. *Documentos de Trabajo N239. Serie Economia*, 2009.

E. Patacchini, E. Rainone, and Y. Zenou. Heterogeneous peer effects in education. *Journal of Economic Behavior & Organization*, 134:190–227, 2017.

D. Powell. Quantile Treatment Effects in the Presence of Covariates. *Quantile Regression*, 2017.

B. Sacerdote. Peer effects with random assignment: Results for Darthmouth roommates. *Quarterly Journal of Economics*, 116(2):681–704, 2001.

B. Sacerdote. Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far? *Handbook of Economics of Education*, 3:249–277, 2011.

B. Sacerdote. Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps Forward? *The Annual Review of Economics*, 1:253–272, 2014.

M. Tincani. Heterogeneous Peer Effects and Rank Concerns: Theory and Evidence. *Working Paper*, 2017.

J.P. Valenzuela, C. Bellei, and D. De los Rios. Socioeconomic school segregation in a market-oriented educational system. The case of Chile. *Journal of Educational Policy*, 104:217–241, 2014.

J. Vigdor and T. Nechyba. Peer effects in North Carolina public schools. *Schools and the Equal Opportunity Problem*, pages 73–101, 2007.

D. Zimmerman. Peer Effects in Academic Outcomes: Evidence from a Natural Experiment. *The Review of Economics and Statistics*, 85(1):9–23, 2003.

A Appendix

A.1 Tables

Table 14: Peer mean and dispersion, OLS and IV (4th grade, Mathematics)

	(1) OLS	(2) OLS	(3) OLS	(4) IV
Peer score Maths	0.420***	0.218***	0.216***	0.160**
	(0.004)	(0.009)	(0.009)	(0.070)
Class SD in maths scores		-0.010***	-0.010***	-0.013**
		(0.004)	(0.004)	(0.006)
Individual controls	Yes	Yes	Yes	Yes
Class level controls	No	Yes	Yes	Yes
School FE	No	Yes	Yes	Yes
Family background controls	No	No	Yes	Yes
Observations	107095	107095	107095	107095
R^2	0.519	0.580	0.582	N/A

Standard errors in parentheses clustered at the class level. Column 1 includes individual measures of previous academic performance measured by the teacher. Column 2 adds class size, class size squared and school fixed effects. Column 3 adds characteristics from the family background including parents' education, books available at home and household monthly income. Column 4 includes the same controls as column 3 but shows the results of the IV estimation, using the shared of displaced students in the class as an instrument for mean ability. F statistic of weak identification for IV: 55

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 15: Peer mean and dispersion, OLS and IV (10th grade, Mathematics)

	(1) OLS	(2) OLS	(3) OLS	(4) IV
Peer score Maths	0.584***	0.263***	0.254***	0.206***
	(0.003)	(0.009)	(0.009)	(0.051)
Class SD in maths scores		-0.017***	-0.018***	-0.019***
		(0.003)	(0.003)	(0.003)
Individual controls	Yes	Yes	Yes	Yes
Class level controls	No	Yes	Yes	Yes
School FE	No	Yes	Yes	Yes
Family background controls	No	No	Yes	Yes
Observations	89788	89788	89788	89788
R^2	0.638	0.672	0.674	N/A

Standard errors in parentheses clustered at the class level. Column 1 includes individual measures of previous academic performance measured by the teacher. Column 2 adds class size, class size squared and school fixed effects. Column 3 adds characteristics from the family background including parents' education, books available at home and household monthly income. Column 4 includes the same controls as column 3 but shows the results of the IV estimation, using the shared of displaced students in the class as an instrument for mean ability. F statistic of weak identification for IV: 64

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 16: Peer mean and dispersion, OLS and IV (4th grade, Spanish)

	(1) OLS	(2) OLS	(3) OLS	(4) IV
Peer score Spanish	0.326***	0.075***	0.073***	0.145**
	(0.004)	(0.010)	(0.010)	(0.058)
Class SD in Spanish scores		-0.081***	-0.081***	-0.062***
		(0.005)	(0.005)	(0.016)
Individual controls	Yes	Yes	Yes	Yes
Class level controls	No	Yes	Yes	Yes
School FE	No	Yes	Yes	Yes
Family background controls	No	No	Yes	Yes
Observations	107095	107095	107095	107095
R^2	0.386	0.438	0.440	N/A

Standard errors in parentheses clustered at the class level. Column 1 includes individual measures of previous academic performance measured by the teacher. Column 2 adds class size, class size squared and school fixed effects. Column 3 adds characteristics from the family background including parents' education, books available at home and household monthly income. Column 4 includes the same controls as column 3 but shows the results of the IV estimation, using the shared of displaced students in the class as an instrument for mean ability. F statistic of weak identification for IV: 65

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 17: Peer mean and dispersion, OLS and IV (10th grade, Spanish)

	(1) OLS	(2) OLS	(3) OLS	(4) IV
Peer score Spanish	0.485***	0.179***	0.168***	0.109
	(0.003)	(0.008)	(0.008)	(0.069)
Class SD in Spanish scores		-0.011***	-0.012***	-0.014***
		(0.003)	(0.003)	(0.004)
Individual controls	Yes	Yes	Yes	Yes
Class level controls	No	Yes	Yes	Yes
School FE	No	Yes	Yes	Yes
Family background controls	No	No	Yes	Yes
Observations	89788	89788	89788	89788
R^2	0.511	0.544	0.548	N/A

Standard errors in parentheses clustered at the class level. Column 1 includes individual measures of previous academic performance measured by the teacher. Column 2 adds class size, class size squared and school fixed effects. Column 3 adds characteristics from the family background including parents' education, books available at home and household monthly income. Column 4 includes the same controls as column 3 but shows the results of the IV estimation, using the shared of displaced students in the class as an instrument for mean ability. F statistic of weak identification for IV: 41

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 18: Peer effects and growth in class size, OLS and IV (4th grade, Mathematics)

	(1) OLS	(2) OLS	(3) OLS	(4) IV
Peer score Maths	0.421***	0.219***	0.217***	0.067
	(0.004)	(0.009)	(0.009)	(0.093)
Growth class size 2009- 2010		-0.001***	-0.001***	-0.002***
		(0.000)	(0.000)	(0.001)
Individual controls	Yes	Yes	Yes	Yes
Class level controls	No	Yes	Yes	Yes
School FE	No	Yes	Yes	Yes
Family background controls	No	No	Yes	Yes
Observations	105674	105674	105674	105674
R^2	0.519	0.580	0.582	N/A

Standard errors in parentheses clustered at the class level. Column 1 includes individual measures of previous academic performance measured by the teacher. Column 2 adds class size, class size squared and school fixed effects. Column 3 adds characteristics from the family background including parents' education, books available at home and household monthly income. Column 4 includes the same controls as column 3 but shows the results of the IV estimation, using the shared of displaced students in the class as an instrument for mean ability. F statistic of weak identification for IV: 37

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 19: Peer effects and growth in class size, OLS and IV (10th grade, Mathematics)

	(1) OLS	(2) OLS	(3) OLS	(4) IV
Peer score Maths	0.585***	0.265***	0.256***	0.208***
	(0.003)	(0.009)	(0.009)	(0.062)
Growth class size 2009- 2010		-0.000	-0.000	-0.001
		(0.000)	(0.000)	(0.000)
Individual controls	Yes	Yes	Yes	Yes
Class level controls	No	Yes	Yes	Yes
School FE	No	Yes	Yes	Yes
Family background controls	No	No	Yes	Yes
Observations	87390	87390	87390	87390
R^2	0.638	0.672	0.674	N/A

Standard errors in parentheses clustered at the class level. Column 1 includes individual measures of previous academic performance measured by the teacher. Column 2 adds class size, class size squared and school fixed effects. Column 3 adds characteristics from the family background including parents' education, books available at home and household monthly income. Column 4 includes the same controls as column 3 but shows the results of the IV estimation, using the shared of displaced students in the class as an instrument for mean ability. F statistic of weak identification for IV: 43

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 20: Peer effects and growth in class size, OLS and IV (4th grade, Spanish)

	(1) OLS	(2) OLS	(3) OLS	(4) IV
Peer score Spanish	0.326***	0.125***	0.124***	0.062
	(0.004)	(0.008)	(0.008)	(0.078)
Growth class size 2009- 2010		-0.002***	-0.002***	-0.002***
		(0.000)	(0.000)	(0.001)
Individual controls	Yes	Yes	Yes	Yes
Class level controls	No	Yes	Yes	Yes
School FE	No	Yes	Yes	Yes
Family background controls	No	No	Yes	Yes
Observations	105674	105674	105674	105674
R^2	0.386	0.438	0.440	N/A

Standard errors in parentheses clustered at the class level. Column 1 includes individual measures of previous academic performance measured by the teacher. Column 2 adds class size, class size squared and school fixed effects. Column 3 adds characteristics from the family background including parents' education, books available at home and household monthly income. Column 4 includes the same controls as column 3 but shows the results of the IV estimation, using the shared of displaced students in the class as an instrument for mean ability. F statistic of weak identification for IV: 37

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 21: Peer effects and growth in class size, OLS and IV (10th grade, Spanish)

	(1) OLS	(2) OLS	(3) OLS	(4) IV
Peer score Spanish	0.487***	0.175***	0.165***	0.039
	(0.003)	(0.008)	(0.008)	(0.095)
Growth class size 2009- 2010		-0.002***	-0.001***	-0.002***
		(0.000)	(0.000)	(0.001)
Individual controls	Yes	Yes	Yes	Yes
Class level controls	No	Yes	Yes	Yes
School FE	No	Yes	Yes	Yes
Family background controls	No	No	Yes	Yes
Observations	87390	87390	87390	87390
R^2	0.511	0.544	0.548	N/A

Standard errors in parentheses clustered at the class level. Column 1 includes individual measures of previous academic performance measured by the teacher. Column 2 adds class size, class size squared and school fixed effects. Column 3 adds characteristics from the family background including parents' education, books available at home and household monthly income. Column 4 includes the same controls as column 3 but shows the results of the IV estimation, using the shared of displaced students in the class as an instrument for mean ability. F statistic of weak identification for IV: 25

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 22: First stages for all grades and subjects from baseline results

	(1) M4	(2) S4	(3) M10	(4) S10
Share of displaced students in the class	-0.923***	-1.098***	-1.059***	-0.992***
	(0.131)	(0.153)	(0.131)	(0.151)
Female	0.007***	0.005*	-0.004	0.003
	(0.002)	(0.002)	(0.003)	(0.003)
GPA 1st grade	0.013***	0.013***		
	(0.003)	(0.003)		
GPA 2nd grade	-0.006	-0.007		
	(0.004)	(0.005)		
GPA 3rd grade	-0.002	-0.003		
	(0.004)	(0.004)		
GPA 7th grade			0.013***	0.012***
			(0.002)	(0.003)
GPA 8th grade			0.037***	0.049***
			(0.004)	(0.004)
GPA 9th grade			0.042***	0.043***
			(0.003)	(0.004)
Class size	0.095***	0.085***	0.007	0.010
	(0.014)	(0.016)	(0.010)	(0.010)
Class size squared	-0.001***	-0.001***	0.000**	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
More than 10 books at home	0.003	0.005**	0.014***	0.016***
	(0.002)	(0.002)	(0.003)	(0.003)
Father education	0.001	0.001*	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Mother education	0.002***	0.001***	0.003***	0.003***
	(0.000)	(0.000)	(0.001)	(0.001)
Household monthly income	0.002***	0.002***	0.004***	0.004***
	(0.000)	(0.001)	(0.001)	(0.001)
Observations	107095	107095	89788	89788

Standard errors in parentheses clustered at the class level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 23: First stages for school selection robustness

	(1) M4	(2) S4	(3) M10	(4) S10
Share of displaced students in the class	-0.463***	-0.590***	-0.988***	-0.954***
	(0.133)	(0.157)	(0.139)	(0.162)
Female	0.006***	0.004	-0.006*	0.002
	(0.002)	(0.003)	(0.003)	(0.003)
GPA 1st grade	0.006**	0.006*		
	(0.003)	(0.003)		
GPA 2nd grade	-0.005	-0.010**		
	(0.004)	(0.005)		
GPA 3rd grade	-0.004	-0.004		
	(0.004)	(0.004)		
GPA 7th grade			0.014***	0.012***
			(0.002)	(0.003)
GPA 8th grade			0.039***	0.051***
			(0.004)	(0.004)
GPA 9th grade			0.043***	0.045***
			(0.004)	(0.004)
Class size	0.048***	0.024	0.007	0.008
	(0.015)	(0.016)	(0.010)	(0.011)
Class size squared	-0.000*	-0.000	0.000**	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
More than 10 books at home	0.000	0.002	0.014***	0.017***
	(0.002)	(0.002)	(0.003)	(0.004)
Father education	0.001	0.000	0.002***	0.003***
	(0.000)	(0.001)	(0.001)	(0.001)
Mother education	0.002***	0.001**	0.004***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
Household monthly income	0.001**	0.001**	0.003***	0.003***
	(0.000)	(0.001)	(0.001)	(0.001)
Observations	87028	87028	81030	81030
R^2	0.908	0.877	0.910	0.888

Standard errors in parentheses clustered at the class level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 24: First stages for all grades and subjects from baseline results with mean and SD of ability.

	(1) M4	(2) S4	(3) M10	(4) S10
Share of displaced students in the class	-0.966***	-1.138***	-1.033***	-0.959***
	(0.130)	(0.152)	(0.130)	(0.150)
Class SD in maths scores	-0.058***	-0.055***		
	(0.009)	(0.010)		
Class SD in Spanish scores			-0.028***	-0.036***
			(0.007)	(0.008)
Female	0.006***	0.004*	-0.004	0.002
	(0.002)	(0.002)	(0.003)	(0.003)
GPA 1st grade	0.013***	0.013***		
	(0.003)	(0.003)		
GPA 2nd grade	-0.005	-0.006		
	(0.004)	(0.005)		
GPA 3rd grade	-0.002	-0.003		
	(0.004)	(0.004)		
GPA 7th grade			0.013***	0.012***
			(0.002)	(0.002)
GPA 8th grade			0.037***	0.049***
			(0.004)	(0.004)
GPA 9th grade			0.043***	0.043***
			(0.003)	(0.004)
Class size	0.098***	0.088***	0.007	0.011
	(0.015)	(0.016)	(0.010)	(0.010)
Class size squared	-0.001***	-0.001***	0.000**	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
More than 10 books at home	0.004*	0.006**	0.014***	0.017***
	(0.002)	(0.002)	(0.003)	(0.003)
Father education	0.001	0.001*	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Mother education	0.001***	0.001***	0.003***	0.003***
	(0.000)	(0.000)	(0.001)	(0.001)
Household monthly income	0.002***	0.002***	0.004***	0.004***
	(0.000)	(0.001)	(0.001)	(0.001)
Observations	107095	107095	89788	89788
R^2	0.884	0.849	0.911	0.888

Standard errors in parentheses clustered at the class level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 25: First stages for all grades and subjects from baseline results and growth of class size.

	(1) M4	(2) S4	(3) M10	(4) S10
Share of displaced students in the class	-0.853***	-1.004***	-0.889***	-0.785***
	(0.141)	(0.164)	(0.135)	(0.156)
Growth class size 2009-2010	-0.003***	-0.004***	-0.005***	-0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
Female	0.007***	0.005*	-0.004	0.003
	(0.002)	(0.003)	(0.003)	(0.003)
GPA 1st grade	0.013***	0.012***		
	(0.003)	(0.003)		
GPA 2nd grade	-0.007	-0.008		
	(0.004)	(0.005)		
GPA 3rd grade	-0.002	-0.003		
	(0.004)	(0.004)		
GPA 7th grade			0.013***	0.011***
			(0.002)	(0.003)
GPA 8th grade			0.037***	0.048***
			(0.004)	(0.004)
GPA 9th grade			0.042***	0.042***
			(0.003)	(0.004)
Class size	0.088***	0.078***	0.018*	0.025**
	(0.014)	(0.015)	(0.010)	(0.011)
Class size squared	-0.001***	-0.001***	0.000*	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
More than 10 books at home	0.002	0.004*	0.013***	0.016***
	(0.002)	(0.002)	(0.003)	(0.003)
Father education	0.001	0.001*	0.002***	0.002***
	(0.000)	(0.001)	(0.000)	(0.001)
Mother education	0.002***	0.002***	0.004***	0.004***
	(0.000)	(0.001)	(0.001)	(0.001)
Household monthly income	0.002***	0.002***	0.003***	0.003***
	(0.000)	(0.001)	(0.001)	(0.001)
Observations	105833	105833	87396	87396
R^2	0.884	0.848	0.911	0.889

Standard errors in parentheses clustered at the class level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$