Teaching practices and student achievement*

Ana Hidalgo-Cabrillana
UAM

Cristina Lopez-Mayan
UAB

February 28, 2014

Preliminary draft

Abstract

Using data from a Spanish assessment program of fourth-grade pupils, this paper analyzes the relationship between standardized student test scores and the practices and materials used by the teacher in class. We group teaching practices and materials in traditional or modern. As a novelty, we construct those aggregate measures using the answers to the same question provided by the teacher and her students. We identify the effect of teaching practices and materials on student achievement by exploiting the between-class variation of those variables within school. Preliminary results show that, when using students' answers, modern teaching has a sizable negative and statistically significant effect on achievement, while traditional teaching has a positive and smaller impact. In contrast, we do not find significant effects of traditional or modern teaching when using teacher answers. Teacher characteristics are not significantly related to student test scores, with the notable exception of a negative relationship if students have a teacher with a high-quality degree.

JEL classification: I20; I21; J24

Keywords: Education; Student achievement; Teacher quality; Sorting

*Financial support from the Ramon Areces grant is gratefully acknowledged. E-mail: crislopez-maya@gmail.com, hidalgocabrillana@gmail.com

1 Introduction

Accumulated research finds that the level of knowledge acquired by people is an important predictor of both individual outcomes (such as future labor market careers) and aggregated outcomes (such as economic growth).[1] In the education production process, it is widely accepted that teachers matter to increase students' achievement.[2] However, the question about what attributes make a teacher more successful than another in enhancing students' performance has not been settled so far. As Hanushek and Rivkin (2006) point out, previous studies do not find consistent evidence that pupils' achievement is strongly correlated to observable teacher characteristics, such as gender, experience, certification, etc. Among exceptions, Rockoff (2004) and, Rivkin et al. (2005), who find significant effects of teacher experience (although the impact is small and concentrated in the first years), and Dee (2005, 2007) who obtain significant effects of teacher's gender and race. The lack of consistent evidence of observed characteristics contrasts with the general finding that teacher effectiveness, measured by teacher fixed effects, has an important impact on student achievement, as shown by Rockoff (2004), and Rivkin et al. (2005). Since observed teacher characteristics can only explain a relatively small part of overall teacher quality, it seems that most of the variation in teacher effectiveness is related to unobserved factors.

Motivated by this lack of results on teacher characteristics, a recent line of research shifts the focus from observed teacher characteristics to teaching practices, that is, what teachers actually do in the classroom (Schwerdt and Wuppermann (2011), Lavy (2011), Bietenbeck (2013)). These studies show evidence that teaching practices matter for student achievement but the findings are not conclusive, especially to identify the best teaching practices.

Most of the current proposals to reform education advocate a greater use of modern teaching practices, that emphasize the ability to process, synthesize, present information, to work with real-world problem solving, in detriment of a traditional rote-learning style. The empirical evidence on how these different teaching styles affect the acquisition of skills is still scarce.

[1] Regarding the effect of human capital on labor market outcomes, see, for example, Murnane et al. (1995), Keane and Wolpin (1997), Cameron and Heckman (1993, 1998), Lazear (2003). With respect to the association between students' test scores and economic growth, see Hanushek and Kimko (2000) and Hanushek and Woessmann (2012).

[2] See Hanushek and Rivkin (2006) and Hanushek (2006) for a review. Recently, Hanushek (2011) quantifies that an effective teacher is equivalent to advancing knowledge in one academic year. Furthermore, Chetty et al. (2011a) conclude that teachers who raise test scores also increases the future income of the student, the probability of attending college and the neighborhood quality, while it decreases the probability of teenage pregnancy.

However, it is an important question since, as a result of the access to massive information facilitated by the technological revolution, people is no longer paid for what they know, but for how they analyze new information and apply it in unfamiliar situations.

This paper adds to the still sparse evidence on the relationship between what teachers do in the classroom and student outcomes. The objective of our paper is to analyze how different practices and materials used by teachers in their class work affect student achievement. Using student and teacher survey data from a Spanish program on student assessment conducted in 2009, *"La Evaluación General de Diagnóstico"* (EGD2009), we group teaching practices and class materials according to whether they are traditional or modern, and we relate those aggregated measures to standardized student test scores.

EGD2009 is a national assessment program that evaluates fourth grade students in four competencies (mathematics, reading, science, and civic competencies). In this paper we focus on mathematics and reading to make our results comparable with previous literature. Importantly, the program is designed to evaluate all students belonging to the same classroom, and in most schools, it evaluates two different classrooms. This makes possible to link each student with the teacher's class. In addition, EGD2009 collects information on the characteristics of students, teachers, principals and schools through a set of questionnaires. Specifically, the teacher and the student questionnaires provide information on the practices and materials used by a teacher in her class work, and we use this to construct our measures of traditional and modern practices following the taxonomy by Zemelman et al. (2005).

Students in fourth grade of primary education are nine years old. The classes for these young children are organized around a teacher (tutor) who teaches them most of the subjects, including maths and reading, and students have the same classmates for the entire school day. In this context, teaching practices are more likely to be teacher-specific than subject-specific: a teacher will choose a teaching style according to their own teaching preferences and/or skills, and it is plausible that, in broad terms, she uses that style for all subjects.

Our empirical strategy exploits the variation in teaching practices and test scores across classes within a school to identify the effect of teaching practices on student achievement. According to the EGD2009 data, around 30% of the observed variation in teaching practices happens within school, that is, between fourth-grade teachers of the same school.

An important concern of this type of analysis is whether there is within-school sorting, that is, if teachers and students are assigned to classrooms according to some nonrandom rule (for instance, better teachers are assigned to classes with better students). The Spanish schooling

system is not track-based in primary education and, consistent with this, we obtain broad evidence of no systematic assignment of teachers and students with specific characteristics to the same classrooms. In addition, although classes are formed randomly, the teacher may still adapt her teaching style to the class level finally formed. We neither obtain evidence that supports this behavior. Nevertheless, our specification controls for a rich set of teacher variables (including tutorial activities and class climate) and student characteristics. Finally, the effect of the teaching practices may be biased if students attending a school present specific characteristics (between-school sorting). This is likely to happen, since the choice of both the neighborhood and the school where to send children are nonrandom decisions of parents. Similarly, each school may have its own teaching philosophy. Since our estimation strategy is based on within-school variation, it deals with the presence of between-school sorting.

Our paper is closely related to three previous works on teaching practices. Schwerdt and Wuppermann (2011) study the effect of lecture-style teaching versus in-class problem solving on standardized test scores using the TIMSS wave of 2003. Using a between-subject strategy to control for unobserved student traits, they find that a ten percentage point shift from problem solving to lecture-style presentation results in an increase in student achievement of about one percent of a standard deviation. This result stands in contrast to constructivist theories of learning. However, it is in line with Brewer and Goldhaber (1997) who conclude that instruction in small groups and emphasis on problem solving lead to lower student test scores.

Lavy (2011) analyzes the effect of traditional and modern teaching on student achievement in Israel using panel data of pupils in fifth and eighth grade. He uses a within-pupil analysis together with school fixed effects to control for unobserved factors. Therefore, the identification strategy is based on the change in exposure to teaching practices among students attending grades five and eight. Lavy (2011) concludes that traditional and modern practices do not necessarily crowd out each other. In particular, practices that emphasize "instilment of knowledge and comprehension", which are viewed as traditional teaching, have a positive effect on test scores, especially of girls and pupils from low socioeconomic backgrounds. "Analytical and critical skills", which are viewed as modern teaching, have also a high payoff, especially among pupils from educated families.

Bietenbeck (2013) analyzes the effect of traditional and modern teaching practices on maths and science test scores using the TIMSS wave of 2007. He estimates a student fixed-effect model to control for unobserved traits. His identification strategy relies on the different

student exposure to teaching practices in maths and science. He concludes that traditional teaching has a positive effect on overall test scores while modern teaching has a statistically insignificant effect. After splitting overall test scores in the scores corresponding to the cognitive skills knowing, applying, and reasoning, he obtains that modern practices have a positive and significant effect on reasoning, while traditional practices increase knowing and applying skills.

Our work extends beyond those previous papers in the following. First, our study analyzes the impact of teaching practices on test scores of younger students (fourth grade of primary school). A recent line of research shows that early interventions in student achievement are key to improve child's success later in life (see Heckman (2008), and Chetty et al. (2011b)). Second, in contrast to Schwerdt and Wuppermann (2011), Lavy (2011), and Bietenbeck (2013), which only have information on teaching practices through the students' answers, we also have that information reported directly by teachers. Therefore, our measures of what teachers really do in the classroom are more reliable. In addition, the same questions on teaching practices are asked to the students. Thus, we can construct also the traditional and modern indexes using students' answers and estimate their impact on test scores. This allows us to compare those estimates with the results obtained using the teacher's answers, and also, to compare our results with previous findings. Third, in addition to practices and materials used in the class, the teacher's questionnaire contains a rich set of information on observed teacher characteristics, including questions about tutorial activities. We use that information to provide new insights on the effect of those variables on student test scores. Finally, none of the previous studies has analyzed the impact of teacher attributes and practices on student achievement in Spain. Therefore, our work provides evidence on the role of the teacher on Spanish student outcomes. Improving our understanding of this is important to solve the serious problems of dropping out (26.5% in 2011 according to Eurostat) and lack of excellence (as shown by the low performance in PISA) that present the current Spanish educational system.

Our preliminary results show that the effect of traditional or modern practices and materials used by the teacher in class is not statistically significant on student test scores when that information is reported by the teacher. However, when the same information is reported by students, we obtain that modern teaching has a sizable negative and statistically significant effect on test scores, while traditional teaching has a smaller and positive impact. Teacher characteristics are not significantly related to student test scores, with the notable exception

of a negative relationship if students have a teacher with a university degree of five years or more. We interpret this effect as evidence that those teachers are negatively selected to teach in primary education.

The rest of the paper is organized as follows. Section 2 describes the database and explains the construction of the teaching practices measures. Section 3 provides a preliminary evidence about the importance of the class level on student test scores. Section 4 discusses the problem of sorting. Section 5 explains the empirical strategy. Section 6 shows the results and Section 7 presents some robustness checks. Finally, Section 8 concludes.

2 Data

We use data from *"La Evaluación General de Diagnóstico"*, a national assessment program conducted for the first time in 2009 by the Instituto Nacional de Evaluación Educativa (INEE), a Spanish institution belonging to the Ministry of Education. This program evaluates four competencies (reading, mathematics, sciences, and civic competencies) to fourth-grade students in primary education using a standardized test.[3] In our analysis, we focus on maths and reading because these are the competencies more often analyzed by researchers. This exam does not have academic consequences for the children. It is only intended to inform children, families and schools about the competencies acquired by students.

EGD2009 evaluates 28,708 pupils belonging to 900 schools following a two-stage stratified sampling design. In the first stage, schools are selected with probabilities proportional to their fourth grade enrollment. In the second stage, one or two fourth grade classes of the school are randomly sampled and all students belonging to these classrooms are evaluated. The sample is designed so that the assessment results are representative at the national and regional level, and by type of school (public/private).

The test consists of both multiple-choice questions and constructed-response items, where the latter requires that students generate and write their own answers. Those type of questions are intended to measure facts, analytical skills and critical thinking.[4] The total student's test score is made available through five *plausible values*, which are values imputed from an estimated proficiency distribution based on the students' answers to the test items. The scores are constructed to have mean equal to 500 and standard deviation equal to 100. We standardize test scores in maths and reading with mean zero and standard deviation one. This

[3]The INEE designed the tests, the sampling and the questionnaires following the PISA design.

[4]For details, see "Evaluación General de Diagnóstico 2009. Informe de resultados".

allows interpreting coefficients as fractions of a standard deviation.

In addition to assess student achievement, EGD2009 collects detailed information through questionnaires filled in by students, families, teachers, and school principals. Fourth grade students are organized into classrooms where a teacher (tutor) teaches them most of the subjects, including the core ones (maths and reading). In Spain, it is usual that students are assigned to a classroom in first grade and they continue with the same classmates until the end of primary education (sixth grade). Pupils also have the same classmates for the entire school day. Therefore, the teacher questionnaire is answered by the tutor of fourth-grade students, who is also their teacher in most of the subjects.

Student and family information includes gender, date of birth, country of origin, household composition, age at starting school, parents' education, parents' labor status, home resources, parents' support in doing homework, parents' expectations about their children, and whether the student repeated.[5] Principal questionnaire includes questions about her personal characteristics (gender, experience, training) and about the school, such as the type of school (public, private and charter), size of the municipality, education levels present at school, rule to group students into classrooms, teaching hours per week in fourth grade, student and parents' involvement, and questions about management duties. The teacher questionnaire provides rich information on personal characteristics (gender, experience, degree, training), the practices and materials used in her class work, subjects taught, tutorial activities and class climate.

The original sample contains 28,708 pupils distributed into 1,358 classrooms in 900 schools. From this initial sample, we drop (i) students with missing maths or reading scores; (ii) classrooms with less than five pupils; (iii) students and teachers with blank questionnaires; (iv) teachers who do not teach maths nor reading, so we are sure that teachers in the final sample teach the subjects evaluated in EGD2009; (v) students and teachers with missing information in basic observed variables (gender, country of origin, parents' education and labor status, household composition, experience, quality of degree)[6]; (vi) teachers with missing information on the items used to construct the teaching practices measures. In addition, in order to deal with the between-school sorting, we drop the schools with only one fourth-grade classroom surveyed. The final sample contains 12,547 students from 763 classrooms and 389 schools.

[5]Regarding household composition we construct two categories: living in single-parent household, and living with siblings. Regarding parents' education, we distinguish the following categories for both parents: primary or less, compulsory, high school, vocational training, and university. Regarding parents' labor status, we construct the following categories: self-employed, employee, unemployed, and inactive.

[6]Since we do control for both parents' education and labor status, we do not use information on home resources to avoid dropping too many individuals from the initial sample.

We have checked that the characteristics of this sample are not significantly different from those in the initial sample (see Appendix A for details). Therefore, the final sample is still representative of the target population of fourth-grade students in Spain.

Table 1 present basic summary statistics describing the fourth grade teachers in primary school in Spain. Fourth grade teachers are mainly women, with more than thirty years of experience, teaching the core subjects of mathematics and reading in classes of an average size of sixteen students.[7] In addition, 74% of teachers teach to the same group of students in third and fourth grade. 17% of teachers present a level of education corresponding to a university degree of five years or a master degree. The rest of teachers hold three-years degree, which is the education level required by law to teach in primary education. Many teachers respond to have participated in some type of training in the last two years. Regarding the work as tutor, teachers meet with parents an average of three time per school year, and it is more usual that the teacher asks for meetings. The characteristics of the learning environment and disciplinary climate are captured by the proportion of warning letters about student's behavior sent to her family, and by the percentage of warnings about temporary class suspension.

Table 2 reports descriptive statistics on observable student characteristics and family background. Around half of fourth-grade students are girls and 5% has repeated at least once. 7% live in single-parent households and most students live with at least one sibling. The proportion of non-Spanish students is 7% and most of them come from Morocco or Latin America. A high percentage of students started school with three years old or less, which is the usual age to start school in Spain. The schooling attainment of mothers and fathers is similar, while the proportion of unemployed or inactive mothers is higher than the proportion of fathers.

Table 3 presents average reading and mathematics test scores. For the full sample, average scores are similar in maths and reading. However, there are differences by gender: on average, girls perform better than boys in reading, while boys perform better in maths. By type of school, average scores are larger in both subjects for students from private schools.

2.1 Teaching practices and materials

The EGD2009 database contains rich information on the teacher who teaches maths and/or reading to all students of a classroom. As commented before, in addition to personal characteristics, the teacher questionnaire provides information on the practices and materials she uses in her class work. In addition, the same questions about teaching practices and materials

[7] Class size is the total number of surveyed students in a classroom.

are asked to the students.

The information related to teaching practices is derived from question 21 of the teacher questionnaire. In this question, teachers are asked "How often do you use the following teaching practices in your lessons this school year?". On a point-four scale with 1 corresponding to "Never or almost never", 2 to "Sometimes", 3 to "Almost always", and 4 to "Always", teachers answer about the following items: (a) "I mostly teach by telling", (b) "Students present works or topics to classmates", (c) "While I teach, I ask questions about the lesson to students", (d) "While I teach, students ask me questions about the lesson", (e) "I promote discussions", (f) "Students work on exercises and activities designed by me", (g) "Students work individually", (h) "Students work in small groups", (i) "I give different exercises or activities to best/worst students". Following the taxonomy by Zemelman et al. (2005), we classify items as reflecting modern or traditional practices. To construct our measures of teaching practices, we exclude item (i) because it reflects the level of students in class and it would leave to a problem of reverse causality in the estimation. We also exclude items (c) and (d) since, according to the teaching taxonomy, it is not possible to unambiguously match those practices as traditional or modern.[8] The classification of the items as traditional or modern teaching practices is displayed in Table 4.

The aggregate measure of traditional teaching practices is the mean of the teacher's answers to items (a), (f) and (g) (see left hand side of Table 4); and the aggregate measure of modern teaching practices is the mean of the teacher's answers to items (b), (e) and (h) (see right hand side of Table 4).

The information related to teaching materials is derived from question 22 of the teacher questionnaire. In this question, teachers are asked "How often do you use in your lessons the following materials?". On the same point-four scale as in question 21, teachers answer about these items: (a) textbook, (b) workbook to do exercises, (c) books from school library, (d) your own materials, (e) newspapers, (f) computers and internet, (g) audiovisual materials. The traditional index is constructed by averaging the teacher's answers to items (a) and (b), and the modern index is constructed as the mean to items (f) and (g).

We use these four indexes (traditional and modern for teaching practices and materials) to estimate the effect on student test scores of what teachers do in the classroom. Unlike previous literature, which uses the information on teaching practices provided by students, we

[8] An additional reason to exclude items (c) and (d) is to minimize the number of observations dropped due to missing information. Nevertheless, we make robustness checks to assess sensitivity of results to include those items.

use the information provided directly by the teacher. Therefore, our measures of the teaching practices are more reliable than those used in previous papers since our measures are reported by the teacher, and so it is likely that our indexes measure more accurately what teachers really do in the classroom. Moreover, the EGD2009 survey also asks students about the same items for teaching practices and materials. In particular, the question on teaching practices is "In general, how is in-class work?". The items to answer about correspond exactly with items (a) to (h) from the teacher questionnaire and they are also coded using the same scale. The question on teaching materials is "How often do you use in the lessons the following materials?", and the items included in the answer are the same as those included in the teacher's question with the exception of item (d).

Therefore, using the same classification of items, we construct modern and traditional indexes of teaching practices and materials by averaging the students' responses at the class level. The results obtained using these measures are useful to compare with results from previous literature, and more importantly, to compare with results obtained using the teacher's answers. Regarding this, it should be noted that the student's question is about all class work and, although most of the subjects are taught by the tutor, the students' answers may refer to all teachers they have in fourth grade.

3 Class effectiveness: a first evidence

In this Section we investigate the importance of teachers and classmates on student achievement. As a previous step to analyze the effect of teaching practices, we want to understand the overall effectiveness of the teacher and the classroom on fourth-grade student achievement in Spain. Previous literature finds that teacher effectiveness, measured by teacher fixed effects, has an important impact on student test scores (Rockoff (2004), and Rivkin et al. (2005)). Moreover, Hanushek (1971) and Jepsen (2005) show that the combination of teachers and classmates, measured by classroom fixed effects, also has a significant impact. For Spain, Anghel and Cabrales (2010) analyze the effectiveness of schools to explain student achievement of sixth-grade pupils in the region of Madrid.

To this end, first, in Table 5 we analyze to what extent, teacher characteristics, class and school explain student variation in test scores of maths and reading. This Table reports R^2 values from a series of regressions where the dependent variable is the student test score in reading (columns 1 to 5) and mathematics (columns 6 to 10). The first specification for each subject includes only student characteristics. The next column adds measured teacher

characteristics (gender, experience, type of degree, taught subjects, training, teacher at 3^{rd} and 4^{th} grade, tutor activities) and class size. The third column for each subject substitutes observed teacher and class variables by classroom fixed effects. The fourth column substitutes the classroom effects by school fixed effects. And the fifth column adds teacher and class variables to the school-fixed effect specification. The results demonstrate quite clearly that the observable class and teacher characteristics explain little of the between-classroom variation in achievement. Importantly, the inclusion of school-fixed effects reduces the explanatory power compared to including only class-fixed effects. These results suggest that much of the variation in student test scores is produced within school, at the class level.

Next, we analyze to what extent the overall teacher and class effectiveness is related to a broad set of observable characteristics.[9] To do this, for each subject, we run a regression of the standardized student test score on student and family characteristics, and on a set of classroom-fixed effects (see equation 1).

$$y_{ics} = \alpha + C_{cs} + \lambda' X_{ics} + \epsilon_{ics} \tag{1}$$

where y_{ics} is the standardized test score of student i in classroom c at school s, C_{cs} is the class fixed effect, and X_{ics} is a vector of student characteristics. In this approach, the class fixed effect is interpreted as the value added or effectiveness of the teacher and the class.

Results are presented in Table 6. Consistent with previous literature, we find that girls perform better in reading than boys, and boys perform better in mathematics than girls: on average, girls get about 0.11 standard deviations more in reading, and boys get 0.13 standard deviations more in mathematics than girls. The age at which a child starts to go to school appears to matter as well: children who start education older than two years old have worse test scores. Indeed, students starting school at six years old get nearly half of a standard deviation less (i.e. 0.47) in reading and 0.29 standard deviation less in mathematics. After controlling for parental education and occupation, on average Latin American and Non-Western Europe students do worse than Spanish children (0.13-0.19 standard deviation in mathematics and 0.21-0.23 standard deviation in reading). No differences appear between Spanish and Moroccan students. Pupils from Asia do worse in reading (0.62 standard deviation less), while students from Western Europe do better than Spanish children in mathematics (0.28 standard deviations more). Parents' education are revealed as an important predictor of student

[9]Note that since each teacher is observed with one classroom, we cannot estimate separately the effect of the teacher from the effect of the students (peer effect).

achievement in both subjects, with the effects ranging between 0.10 to 0.40 standard deviations. The impact is higher than the labor status of the parents. Indeed, we only find significant effects for the mother's labor status. Interestingly, students whose mothers are employee instead of self-employed perform better (0.07 standard deviation more). We do not observe significant differences for children with inactive or unemployed mothers (except for inactive mothers in maths, although the effect is only significant at 10% level). Living in a single-parent household does not have a significant effect while living with siblings decrease reading test score by 0.06 standard deviation. We find that having repeated a grade or being born in the fourth quarter predict a negative effect on test scores in both subjects (around 0.2-0.3 standard deviations less for repeater and 0.15-0.17 standard deviations less for being born in the fourth quarter).

After obtaining the overall class effectiveness once differences in students characteristics are controlled for, we run a second-stage regression of the estimated class fixed effect on different sets of observed teacher, principal and school characteristics, school-fixed effects and class level variables. We also include the indexes of traditional and modern teaching practices reported by teachers.

$$C_{cs} = \gamma\, TP_{cs} + \lambda' T_{cs} + \beta' X_{cs} + S_s + v_{cs} \tag{2}$$

where C_{cs} is the estimated class effect from first-stage regression, TP_{cs} is the vector of teaching practices in class (teacher) c of school s, T_{cs} is a vector of other teacher characteristics, X_{cs} is a vector of student characteristics measured at the class level, and S_s is a school-fixed effect. In some specifications, we substitute the school-fixed effect by a vector of school and principal characteristics. Because of differences in the size across classrooms, we estimate equation 2 weighted by the inverse of the estimated standard error of C_{cs}. The objective of estimating this equation is to relate the overall class effectiveness with observed class and teacher characteristics, once we have controlled for the school.

Tables 7 and 8 contain the results of the second-stage estimations for mathematics and reading, respectively. Because of the endogeneity of most of the control variables, we gradually add variables in our estimations in order to test the sensitivity of the coefficients. Column (1) includes only the teaching practices indexes. Column (2) adds school-fixed effects, and column (3) substitutes the fixed effects by school and principal observed characteristics. Column (4) adds to the first specification other teacher and class characteristics. Column (5) includes teacher, school and principal variables. Column (6) includes teacher and class characteristics and school fixed effects. Finally, column (7) substitutes school fixed effect by school and

principal variables.

In maths, we obtain that modern teaching practices are negatively correlated with class effectiveness across specifications, while the relationship with traditional teaching practices is positive. However, the estimated effects are small and imprecisely measured in most specifications. In reading, there is no clear pattern in the relationship of traditional and modern practices with class effectiveness. With respect to other teacher characteristics, we find that the years of experience are significantly associated with class effectiveness, although the effects vary by subject: in maths, the relationship is negative and significant for teachers with five to nine years of experience, and in reading, the relationship is positive and significant for highly experienced teachers. Interestingly, teachers who taught the same class in third grade are associated with higher effectiveness in fourth grade, although the effect is only significant in maths. In addition, the higher the proportion of warnings sent to the family, the lower the class effectiveness. This may reflect that a bad class climate is negative for learning, although the effect is only significant for learning maths.

Regarding class-level variables, we find that the higher class size, the higher class effectiveness, although the effects are quite small. Once we include school-fixed effects, students' characteristics measured at class level are mostly non-significant, reflecting that family characteristics matter only across schools (between-school sorting). Among the school variables, the number of teaching hours per week is significant and positively related to class effectiveness in both subjects. Principals who received training in the last years are negatively associated with class effectiveness, while those who are female of have a five-years degree present a positive relationship with effectiveness in reading.

The evidence from this second-stage analysis seems to suggest, like previous papers, that observed teacher characteristics are not strongly associated with class effectiveness. We neither find strong evidence of the relationship between teaching practices and class effectiveness. However, this preliminary evidence needs to be deeply analyzed.

4 Selection of students and teachers to classes

In this Section we discuss the assignment of students between and within schools. An important concern of the analysis is whether the teaching style depends on the characteristics of the students in the class. The existence of any type of sorting will bias the estimates of the effect of the teaching practices due to the correlation of this variable with unobserved teacher and student characteristics.

One possible type of sorting is between-school sorting. It arises, for instance, if students attending a school present specific characteristics. This is likely to happen, since the choice of both the neighborhood and the school where to send children are nonrandom decisions of parents. Between-school sorting can also arise if a school hires systematically teachers with some specific characteristics or if it has a certain teaching philosophy. We deal with the presence of between-school sorting by focusing only on the schools with two sampled classrooms. Therefore, the identification of the effect of the teaching practices will rely on within-school variation of teacher's practices and student's test score. In this way, we control for the endogenous selection of students and teachers across schools.

A second type of sorting occurs within school. This within-school sorting may arise if teachers and students are assigned to classrooms according to some nonrandom rule. This happens, for instance, if parents try to influence on the teacher assigned to their children ("teacher shopping"). However, in Spain, this is not a concern, since parents cannot influence who is the teacher assigned to a classroom. Within-school sorting may also appear if the school principal assigns students to teachers following a nonrandom rule (for instance, if better teachers are assigned to classes with better students). The random assignment of students and teachers to classrooms avoids the bias coming from the within-school sorting by breaking the link between teaching practices and class-level characteristics. However, the random assignment of students to classes is not a common rule in any country, and, therefore, within-school sorting is a potential concern in the analysis of student achievement for most countries, specially for those where the schooling system is strongly track-based. Although this is not the case in Spain, especially in primary education, we conduct several types of analysis to shed light on the presence of within-school sorting among fourth-grade students.

In Table 9, we present the answers to a question about the rule to group students. This question is included in the principal's questionnaire, who is the responsible for the assignment of students to classes. 44% of principals group students following a random criteria, and 77% of the principals respond that they form classes so that they are well balanced, mixing students with different characteristics. Only 11% of principals mention grouping students by academic ability.[10] Therefore, most of the principals report that they group fourth-grade students more or less randomly.

This institutional evidence about the random formation of classes within schools is complemented with a series of regression analysis. First, we investigate whether students with

[10]Note that this question admits multiple answers.

certain family characteristics are more likely to be in classes with a certain type of teacher. To this end, we regress different observed teacher variables on the sociodemographic characteristics of the classroom.[11] These characteristics are constructed as the class-level average of individual student characteristics. Tables 12 and 13 report the results from these regressions. The main conclusion is that there is no systematic assignment of particular teachers to classes with particular characteristics (i.e. high proportion of immigrants, repeaters, students from low-educated families, etc).

Second, we investigate whether classrooms that differ in teaching practices, differ in pupils' characteristics as well. For this purpose, following Lavy (2011), we run a set of regressions of student level characteristics on the teaching practices indexes reported by the tutor and school fixed effects (balancing tests).

$$x_{ics} = \alpha + S_s + \gamma\, TP_{cs} + v_{ics} \tag{3}$$

Table 10, which presents the results from these regressions, shows little evidence in favor that students of different family backgrounds are more likely to be assigned to teachers using certain teaching practices, conditional on the school they attend.

Therefore, previous evidence suggests that, within school, there is no systematic assignment of students with certain characteristics to certain teachers who use certain teaching practices. However, even though classes are formed more or less randomly, they may receive other school resources differentially. For instance, a teacher with a certain teaching style may be assigned to classes of certain size. To check this, we run regressions of the teaching practices on class size and class size squared. Results are shown in Table 11. For the modern and traditional teaching indexes constructed with teacher's answers, we run two regressions: without including school fixed effects (columns 1 and 3), and adding school fixed effects (columns 2 and 4). Results show that there is no systematic correlation between class size and the two teaching practices indexes. All estimates are not significantly different from zero.

5 Empirical Strategy

Using observational data to estimate the causal effect of the teaching practices on student test scores has to deal with the potential problem of within-school sorting. The evidence shown in previous Section suggests that in the EGD2009 data there is non-systematic assignment of

[11] See, for instance, Jepsen (2005).

students to teachers within school. Nevertheless, the presence of within-school sorting may still be a problem. Thus, we adopt an empirical strategy that addresses this possible concern.

Fourth-grade students in Spain are organized into classes with the same classmates for the entire school day. Usually, they also have the same classmates along all grades of primary education. Each class is assigned a teacher (tutor) who teaches them most subjects, and who makes tutorial work, such as meeting with students' parents to talk about the achievement of their children or existing class-disruptive problems. Frequently, the tutor teaches the core subjects (maths and reading). In EGD2009 data, we can observe this for 88% of tutors (see Table 1). In this context, teaching practices are more likely to be teacher-specific than subject-specific: a teacher will choose a teaching style according to their own teaching preferences and/or skills, and it is plausible that, in broad terms, she uses that style for all subjects.[12]

EGD2009 evaluates all students belonging to the same classroom, so this makes possible to link each student with the teacher's class. We observe the teaching practices of each class' tutor and we observe students twice, once in maths and once in reading. Using the matched-pairs data of EGD2009, we adopt an empirical strategy that exploit the variation in teaching practices across classes within a school to identify the effect of teaching practices on student achievement.

Our starting point is a standard education production function:

$$y_{icsj} = \alpha + \gamma' TP_{cs} + \lambda' T_{cs} + \beta' X_{ics} + \phi' S_s + \mu_{icj} + \eta_{cs} + \tau_s + \varepsilon_{icsj} \tag{4}$$

where y_{icsj} is the standardized test score of student i in classroom c at school s in subject $j = \{m, r\}$ (m refers to maths and r to reading). TP_{cs} is the vector of teaching practices in class (teacher) c of school s. T_{cs} is a vector of teacher characteristics and class size. X_{ics} is a vector of student i characteristics. S_s are school characteristics. μ_{icj}, η_{cs} and τ_s are, respectively, student, teacher and school unobserved traits. ε_{icsj} is an idiosyncratic shock. The presence of the unobserved variables would bias the estimate of γ.

Averaging across the students in class c and then across classes of the same school s:

$$\overline{y}_{sj} = \alpha + \gamma' \overline{TP}_s + \lambda' \overline{T}_s + \beta' \overline{X}_s + \phi' S_s + \overline{\mu}_j + \overline{\eta}_s + \tau_s + \overline{\varepsilon}_{sj} \tag{5}$$

We construct the deviation of the student test score to the school average. That is, we

[12] Bietenbeck (2013) assumes instead that teacher practices are subject specific because a different teacher teaches maths and science, and students are asked about the teaching practices in each subject.

take differences of (4) and (5), so the headline specification is:

$$y_{icsj} - \overline{y}_{sj} = \gamma'(TP_{cs} - \overline{TP}_s) + \lambda'(T_{cs} - \overline{T}_s) + \beta'(X_{ics} - \overline{X}_s) + (\mu_{icj} - \overline{\mu}_j) + (\eta_{cs} - \overline{\eta}_s) + (\varepsilon_{icsj} - \overline{\varepsilon}_{sj}) \quad (6)$$

Therefore, for each subject j, the deviation of the student test score to the mean of the test scores across the two classes within the same school is related to the deviation with respect to the school average of the teaching practices, teacher observed characteristics, students' socioeconomic background, and unobserved terms. In particular, unobserved factors affecting differences in individual achievement with respect to the school average achievement in subject j are potentially the following:

$(\mu_{icj} - \overline{\mu}_j)$: the deviation of the unobserved ability of student i in subject j with respect to the school average unobserved ability in that subject.

$(\eta_{cs} - \overline{\eta}_s)$: the deviation of the unobserved teacher ability or motivation in class c with respect to the mean unobserved ability or motivation of the two teachers in the school s.

$(\varepsilon_{icsj} - \overline{\varepsilon}_{sj})$: differences in the idiosyncratic shock.

In this empirical strategy we exploit the variation in teaching practices across classes within a school to identify the effect of teaching practices on student achievement (γ). According to the data, around 30% of the observed variation in teaching practices happens within school (remaining 70% happens between schools). Our identifying assumption is that the deviation of the teaching practices in class c with respect to the average teaching practices within school is uncorrelated with the error term conditional on the other regressors. However, the presence of the terms $(\mu_{icj} - \overline{\mu}_j)$ and $(\eta_{cs} - \overline{\eta}_s)$ may yield to obtain biased estimates of γ. In particular, γ would be biased:

- If the student unobserved ability is different from the school average unobserved ability in subject j ($\mu_{icj} - \overline{\mu}_j \neq 0$), and, at the same time, this difference has a direct effect on $y_{icsj} - \overline{y}_{sj}$ while it is correlated with the teaching practices, that is, if $corr(\mu_{icj} - \overline{\mu}_j, TP_{cs} - \overline{TP}_s) \neq 0$. This would happen if there is sorting of students to classes within school (so, the ability composition of the two classes will be different) and the teacher adapts her teaching practices to the resulting level of ability in the class. For example, if high-ability students are assigned to the same class and the teacher decides to use more modern teaching practices with those students, the estimate of γ will be biased. It is important to note that if $\mu_{icj} - \overline{\mu}_j \neq 0$ but teachers do not adapt their teaching style to the ability level of the class, γ will not be biased.

Regarding this source of bias, in the previous Section we have analyzed the presence of sorting within school and we do not find evidence of systematic assignment of students to classes or teachers. Nevertheless, in our specification we include $X_{ics} - \overline{X}_s$ to control for differences in the student background characteristics across classes.

- If the unobserved teacher ability or motivation is different from the school average unobserved teacher traits, $(\eta_{cs} - \overline{\eta}_s \neq 0)$. The bias would appear if that difference has a direct impact on $y_{icsj} - \overline{y}_{sj}$, while, at the same time, it is correlated with the teaching practices. In other words, if teachers of certain ability sort into certain teaching practices and teacher's ability affects student test scores directly and not only through the teaching practices channel, then γ would be biased.

  We address this omitted-variable problem by controlling for a broad set of observed teacher characteristics, including not only the typical controls used in the literature (gender, experience, degree, training), but also variables capturing teacher's work as tutor. This will minimize the presence of unobserved teacher ability. Although we cannot fully disregard the presence of such unobserved teacher traits, and consequently we refrain from interpreting our estimates as causal, we should note that our variable of teaching practices is potentially less endogenous with respect to the test scores in subject j since the tutor answers about her class teaching practices and not about her particular teaching style in subject j.

6 Results

Tables 14 to 17 present the estimation results of different specifications of equation (6). The dependent variable is the difference between the student test score and the mean of the students test scores from the two classes within the same school. In Table 14 we estimate an specification that includes the indexes of modern and traditional teaching practices. We run separate regressions for the indexes constructed using the teacher's or the students' answers. In Table 15, we repeat this analysis but including the indexes of modern and traditional teaching materials instead of the indexes of teaching practices. In both Tables, results are obtained for the pool of subjects, including a subject dummy. In Tables 16 and 17 we present the results of estimating separately for maths and reading, respectively. For the sake of brevity, we only report the estimated coefficients of the teaching indexes and teacher characteristics, although the rest of coefficients are available upon request.

Columns (1)-(3) of Table 14 contain the results of the estimation when teaching practices are reported by the tutor, and columns (4)-(6) contain the results when they are reported by students. Specification in columns (1) and (4) only includes the teaching practices measures and a dummy for mathematics. In columns (2) and (5) we add class size and teacher characteristics, and in columns (3) and (6), we also include contextual characteristics of students. This last specification corresponds to equation (6). We obtain that the effect of teaching practices on test scores differs depending on who reports the information. Specifically, when it is reported by the tutor, both the effect of traditional and modern practices is small and non statistically significant. However, when the information is reported by the student, the effect of the two types of practices is significant at one percent level. Traditional practices have a positive effect of around 0.06 standard deviations while modern practices reduce test scores by 0.20 standard deviations. The two effects are robust across specifications (see columns (4) to (6)). It is interesting to note that the negative effect of modern practices becomes positive, although non significant, when practices are reported by the tutor.

In line with the previous literature, most teacher characteristics (female, years of experience, training) have a statistically insignificant effect on student achievement. We neither find a significant relationship between test scores and taught subjects, class climate and tutorial work. Regarding the latter, the only exception is the positive effect (0.05 standard deviations) on student achievement if the teacher was the tutor of the class both in third and fourth grade. This effect is only significant in the regressions when teaching practices are reported by the students. The most interesting effect is found for the type of degree that the tutor holds. Teachers who have a university degree of five years or more reduce student achievement by 0.07 standard deviations compared to teachers with a three-years degree. The effect is clearly significant and robust across all specifications. Since holding a three-years degree is the established requirement to teach in primary education in Spain, this negative effect may suggest that teachers with a five-years degree teaching in primary education are negatively self-selected. That is, they enter primary education after failing to find a job in the private sector and/or in secondary education (where the requirement is to hold at least a five-years degree). Consequently, those teachers may lack motivation and good teaching skills, and this would explain the negative effect that we find.

In Table 15 we present the results from the same specifications shown in Table 14 after substituting the teaching practices by the teaching materials. The main conclusion is that results hardly change. Consistent with the evidence obtained when including teaching prac-

tices, we also obtain that teaching materials are not significant when they are reported by the tutor, while they are significant when reported by the pupils. In this case, we also find that using modern (traditional) materials is related to lower (higher) test scores. The significance and magnitude of these effects is the same as in Table 14. The effects of the rest of teacher characteristics are also very similar to those found in Table 14, including the negative effect of holding a degree of five years or more.

Tables 16 and 17 show the effect of teaching practices and materials obtained by estimating separately for maths and reading test scores, respectively. We present the results for the regression that includes all controls (teacher and contextual characteristics). Columns (1) and (2) report the results from the estimation including teaching practices reported by the tutor and the students, respectively. Columns (3) and (4) report the results corresponding to the teaching materials reported by the tutor and the students, respectively. Results are very similar to those obtained by pooling the two subjects. When reported by the tutor, teaching practices and materials are not significant. When reported by the students, traditional teaching is positively related to test scores while modern teaching is negatively related. The only difference with previous results is that modern teaching practices reported by the tutor have a positive effect on reading test scores, significant at ten percent level (see column (1) of Table 17). This contrasts with the negative sign found when the index is constructed using students' answers (column (2) of Table 17). With respect to the rest of regressors, we find that teachers with a degree of five years or more reduce test scores in both maths and reading. Class size has a significant, although negligible, effect but only when students' answers are used to construct the teaching indexes.

All in all, our results show, first, that the type of teaching practices and materials used in class is significantly related to student achievement when this information is reported by the students. Specifically, modern practices and materials have a sizable negative effect on achievement while traditional practices and materials have a positive, smaller, effect. However, when teaching practices and materials are asked to the tutor we do not find any significant effect on student test scores. Therefore, our results provide new insights on the effect of teaching on achievement. The evidence obtained by previous papers is based on teaching practices reported by students. Schwerdt and Wuppermann (2011) find that teachers who spend more time lecturing are associated with higher test scores. Lavy (2011) obtains that both traditional and modern teaching practices have positive effects on test scores, although larger for traditional teaching. Bietenbeck (2013) concludes that only traditional teaching

has a statistically significant and positive effect on overall test scores. However, our evidence suggests that, once we differentiate who reports the information on teaching practices (students or tutor), the results change dramatically.

In addition, our results show that having a teacher with a university degree of five years or more is negatively related to student performance. This result differs from previous literature, which finds that teachers with a higher quality degree are related to better student performance. This may suggest that teachers with a five-years degree teaching in primary education in Spain are negatively selected. This should be taken into account for educational authorities when they design the degree requirements to allow access to teach in primary education.

7 Robustness Checks

TO BE COMPLETED

8 Conclusions

TO BE COMPLETED

References

Anghel, B. and A. Cabrales (2010): "The Determinants of Success in Primary Education in Spain," Documento de Trabajo FEDEA 2010-20.

Bietenbeck, J. C. (2013): "Teaching Practices and Cognitive Skills," Mimeo CEMFI.

Brewer, D. J. and D. D. Goldhaber (1997): "Why don't schools and teachers seem to matter?" *The Journal of Human Resources*, 32, 505–523.

Cameron, S. V. and J. J. Heckman (1993): "The Nonequivalence of High School Equivalents," *Journal of Labor Economics*, 11, 1–47.

——— (1998): "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males," *Journal of Political Economy*, 106, 262–333.

Chetty, R., J. Friedman, and J. Rockoff (2011a): "The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood," NBER Working Paper No. 17699.

Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan (2011b): "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star," *Quarterly Journal of Economics*, 216, 1593–1660.

Dee, T. S. (2005): "A Teacher like Me: Does Race, Ethnicity, or Gender Matter?" *American Economic Review Papers and Proceedings*, 95, 158–165.

——— (2007): "Teachers and the Gender Gaps in Student Achievement," *Journal of Human Resources*, 42, 528–554.

Hanushek, E. (2011): "The Economic Value of Higher Teacher Quality," *Economics of Education Review*, 30, 466–479.

Hanushek, E. A. (1971): "Teacher characteristics and gains in student achievement: estimation using micro data," *American Economic Review*, 61, 280288.

——— (2006): *School resources*, in Handbook of the Economics of Education, edited by Eric A. Hanushek and Finis Welch. Amsterdam: North Holland, vol. 2, chap. 14, 865–908.

Hanushek, E. A. and D. D. Kimko (2000): "Schooling, labor force quality, and the growth of nations," *American Economic Review*, 90, 1184–1208.

Hanushek, E. A. and S. Rivkin (2006): *Teacher Quality*, in Handbook of the Economics of Education, edited by Eric A. Hanushek and Finis Welch. Amsterdam: North Holland, vol. 2, chap. 18, 1051–1078.

Hanushek, E. A. and L. Woessmann (2012): "Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation," *Journal of Economic Growth*, 17, 267–321.

Heckman, J. J. (2008): "Schools, Skills and Synapses," *Economic Inquiry*, 46, 289–324.

Jepsen, C. (2005): "Teacher characteristics and student achievement: Evidence from teacher surveys," *Journal of Urban Economics*, 57, 302–319.

Keane, M. P. and K. I. Wolpin (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105, 473–522.

Lavy, V. (2011): "What Makes an Effective Teacher? Quasi-Experimental Evidence," NBER Working Paper No. 16885.

Lazear, E. P. (2003): "Teacher Incentives," *Swedish Economic Policy Review*, 10, 179–241.

Murnane, R. J., J. B. Willet, and F. Levy (1995): "The Growing Importance of Cognitive Skills in Wage Determination," *Review of Economics and Statistics*, 77, 251–266.

Rivkin, S. G., E. A. Hanushek, and J. F. Kain (2005): "Teachers, Schools, and Academic Achievement," *Econometrica*, 73, 417–458.

Rockoff, J. E. (2004): "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data," *American Economic Review Papers and Proceedings*, 94, 247–252.

Schwerdt, G. and A. C. Wuppermann (2011): "Is traditional teaching really all that bad? A within-student between-subject approach," *Economics of Education Review*, 30, 365–379.

Zemelman, S., H. Daniels, and A. Hyde (2005): *Best Practice. Today's Standards for Teaching and Learning in America's Schools*, Heinemann, third ed.

Tables

Table 1: Descriptive statistics of teacher

	Mean	Std. Dev.	Observations
Female	0.75	0.43	778
Experience (years):			
Less than 5	0.10	0.30	778
5 - 9	0.11	0.31	778
10 - 14	0.06	0.25	778
15 - 19	0.09	0.29	778
20 - 24	0.10	0.30	778
25 - 29	0.16	0.36	778
30 or more	0.39	0.49	778
5-years degree or more	0.17	0.38	778
Class size	16.43	4.58	778
Taught subjects:			
Reading and Maths	0.88	0.32	778
Reading	0.05	0.22	778
Maths	0.07	0.25	778
Training:			
Attending courses	0.71	0.45	778
Working teams at school	0.62	0.49	778
Congresses and teaching projects	0.46	0.50	778
Type of warnings to students:			
Letter to the family	0.65	0.48	778
Temporary class suspension	0.16	0.37	778
Person asking for a meeting:			
Parents	0.21	0.41	774
Teacher	0.34	0.47	774
Number of meetings with students' parents	3.03	0.97	770
Teacher at 3^{rd} and 4^{th} grades	0.74	0.44	771

Table 2: Descriptive statistics of students

	Mean	Std. Dev.
Female	0.49	0.50
Repeater	0.05	0.23
Born in 4th quarter	0.32	0.47
Living in single-parent household	0.07	0.26
Living with siblings	0.84	0.36
Country of origin:		
Spain	0.92	0.27
Western Europe	0.00	0.05
Non-Western Europe	0.02	0.13
Morocco	0.00	0.07
Latin America	0.05	0.21
Asia	0.00	0.05
Other	0.00	0.06
Age at starting school:		
2 years old or less	0.60	0.49
3 years old	0.36	0.48
4 years old	0.03	0.17
5 years old	0.01	0.11
6 years old	0.01	0.07
Mother's education:		
Primary or less	0.10	0.30
Compulsory	0.24	0.43
High School	0.14	0.35
Vocational training	0.20	0.40
University	0.31	0.46
Father's education:		
Primary or less	0.13	0.33
Compulsory	0.26	0.44
High School	0.15	0.36
Vocational training	0.20	0.40
University	0.26	0.44
Mother's labor status:		
Self-employed	0.13	0.34
Employee	0.52	0.50
Unemployed	0.10	0.30
Inactive	0.25	0.43
Father's labor status:		
Self-employed	0.26	0.44
Employee	0.65	0.48
Unemployed	0.07	0.26
Inactive	0.02	0.13
Observations	12785	

Table 3: Average test scores

	Maths	Reading
Full sample	516.10	516.32
By gender:		
Male	523.14	511.32
Female	508.84	521.47
Gap (male-female)	14.30	-10.15
By type of school		
Public	507.06	506.8
Private	532.32	533.42
Gap (public-private)	-25.26	-26.62
Number of individuals	12785	12785
Number of classrooms	778	778
Number of schools	389	389

Table 4: Matched teacher questionnaire items

Traditional Teaching	Modern Teaching
Item (a): I mostly teach by telling	Item (b): Students present works or topics to classmates
Item (f): Students work on exercises and activities designed by me	Item (e): I promote discussions
Item (g): Students work individually	Item (h): Students work in small groups

Teachers respond to the question "How often do you use the following teaching practices in your lessons this school year?". Answers are coded on a point-four scale with 1 corresponding to "Never or almost never", 2 to "Sometimes", 3 to "Almost always", and 4 to "Always".

Table 5: Comparison of explanatory power

	Reading					Maths				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Explanatory variables										
Student characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Teacher characteristics	no	yes	no	no	yes	no	yes	no	no	yes
Class fixed effects	no	no	yes	no	no	no	no	yes	no	no
School fixed effects	no	no	no	yes	yes	no	no	no	yes	yes
R^2	0.136	0.148	0.260	0.227	0.230	0.115	0.131	0.250	0.215	0.217
Observations	12547	12547	12547	12547	12547	12547	12547	12547	12547	12547
Classrooms	763	763	763	763	763	763	763	763	763	763
Schools	389	389	389	389	389	389	389	389	389	389

Dependent variables are standardized maths and reading test scores. Student characteristics: female, country of origin, age at starting school, parents' education, parents' labor status, repeater, living in single-parent household, living with siblings, born in 4th quarter. Teacher characteristics (female, experience, 5-years degree or more, taught subjects, training, type of warnings to students, person asking for a meeting, number of meetings with students' parents, teacher at 3^{rd} and 4^{th} grades) and class size.

Table 6: First stage of Class Fixed Effect Analysis

	Maths	Reading
Student characteristics		
Female	-0.13***	0.11***
	(0.02)	(0.02)
Age at starting school:		
3 years old	-0.02	-0.04**
	(0.02)	(0.02)
4 years old	-0.13**	-0.14**
	(0.05)	(0.06)
5 years old	-0.13	-0.34***
	(0.08)	(0.07)
6 years old	-0.29**	-0.47***
	(0.12)	(0.10)
Country of origin:		
Western Europe	0.28*	-0.37*
	(0.15)	(0.19)
Non-Western Europe	-0.13*	-0.23***
	(0.06)	(0.07)

Dependent variables: standardized maths and reading test scores.
Standard errors clustered at the classroom level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. (*Continued on next page*)

Table 6: (continued)

	Maths	Reading
Student characteristics		
Morocco	-0.04	-0.09
	(0.11)	(0.11)
Latin America	-0.19***	-0.21***
	(0.05)	(0.05)
Asia	-0.20	-0.62***
	(0.20)	(0.19)
Other	-0.32**	-0.34**
	(0.14)	(0.15)
Mother's education:		
Compulsory	0.11***	0.10***
	(0.03)	(0.03)
High School	0.19***	0.22***
	(0.04)	(0.04)
Vocational training	0.18***	0.16***
	(0.04)	(0.04)
University	0.37***	0.36***
	(0.04)	(0.04)

Dependent variables: standardized maths and reading test scores.
Standard errors clustered at the classroom level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. (*Continued on next page*)

Table 6: (continued)

	Maths	Reading
Student characteristics		
Mother's labor status:		
Employee	0.07***	0.07**
	(0.03)	(0.03)
Unemployed	0.03	0.04
	(0.03)	(0.04)
Inactive	0.05*	0.04
	(0.03)	(0.03)
Father's education:		
Compulsory	0.05	0.11***
	(0.03)	(0.03)
High School	0.18***	0.29***
	(0.04)	(0.03)
Vocational training	0.14***	0.20***
	(0.03)	(0.03)
University	0.31***	0.40***
	(0.04)	(0.04)
Father's labor status:		
Employee	-0.00	0.01
	(0.02)	(0.02)
Unemployed	-0.05	-0.01
	(0.04)	(0.04)
Inactive	-0.05	-0.01
	(0.06)	(0.07)
Repeater	-0.21***	-0.28***
	(0.02)	(0.02)
Born in 4th quarter	-0.15***	-0.17***
	(0.02)	(0.02)
Single-parent household	-0.03	-0.06
	(0.03)	(0.03)
Siblings	-0.01	-0.06***
	(0.02)	(0.02)
Observations	12547	12547
R^2	0.25	0.26
Class FE	Yes	Yes

Dependent variables: standardized maths and reading test scores.
Standard errors clustered at the classroom level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Second stage of Class FE Analysis (Maths)

	Dependent variable: Class fixed effect						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Traditional Teaching Practices	0.02	0.02	0.02	0.02	-0.00	0.01	0.02
	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.05)	(0.03)
Modern Teaching Practices	-0.06*	-0.02	-0.05	-0.01	-0.04	-0.01	-0.05*
	(0.04)	(0.05)	(0.04)	(0.05)	(0.03)	(0.05)	(0.03)
School characteristics:							
Public school			0.02		-0.02		0.05
			(0.11)		(0.11)		(0.11)
Size of municipality:							
2000-10000			0.22		0.28		0.29
			(0.18)		(0.19)		(0.18)
10000-50000			0.12		0.17		0.22
			(0.18)		(0.19)		(0.18)
50000-500000			0.14		0.20		0.23
			(0.18)		(0.19)		(0.18)
$>$ 500000			0.26		0.29		0.27
			(0.19)		(0.20)		(0.18)
Education levels at school:							
Preschool			-0.12		-0.11		-0.08
			(0.11)		(0.10)		(0.10)
Primary			0.33		0.13		-0.11
			(0.32)		(0.21)		(0.25)
Compulsory			0.10		0.01		-0.00
			(0.12)		(0.11)		(0.11)
Academic high school			-0.04		-0.03		-0.04
			(0.06)		(0.06)		(0.06)
Vocational high school			0.07		0.08		0.09
			(0.13)		(0.13)		(0.13)
Vocational college			-0.02		-0.09		-0.13
			(0.15)		(0.15)		(0.14)
Teaching hours per week			0.01***		0.01***		0.01***
			(0.00)		(0.00)		(0.00)
Rule to group students:							
Alfabethic			0.04		0.02		0.01
			(0.04)		(0.03)		(0.03)
Gender equilibrium			0.12***		0.07**		0.05

Weighted regressions. Standard errors clustered at the classroom level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. (*Continued on next page*)

Table 7: Second stage of Class FE Analysis (Maths) (continued)

	Dependent variable: Class fixed effect						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			(0.04)		(0.04)		(0.04)
Language			-0.02		-0.02		-0.05
			(0.05)		(0.04)		(0.04)
Academic outcomes			0.07		0.07*		0.07*
			(0.05)		(0.04)		(0.04)
Searching homogeneity			0.03		0.05		0.06
			(0.05)		(0.04)		(0.04)
Searching heterogeneity			0.00		0.00		-0.01
			(0.04)		(0.03)		(0.03)
Others			0.03		0.01		-0.01
			(0.04)		(0.03)		(0.03)
Principals in last 10 years:							
2			0.03		0.02		-0.01
			(0.05)		(0.05)		(0.05)
3			-0.01		-0.01		-0.05
			(0.06)		(0.06)		(0.05)
> 3			0.01		0.00		-0.05
			(0.06)		(0.06)		(0.06)
Principal characteristics:							
Female			0.04		0.05		0.03
			(0.03)		(0.03)		(0.03)
Non-university			0.05		-0.01		0.00
			(0.06)		(0.06)		(0.05)
5-years degree			0.04		0.02		0.01
			(0.03)		(0.03)		(0.03)
Experience (years):							
5-9			0.03		0.05		0.03
			(0.04)		(0.04)		(0.03)
10-14			0.02		0.01		-0.01
			(0.06)		(0.05)		(0.05)
15-19			-0.02		0.01		0.00
			(0.06)		(0.06)		(0.06)
20-24			0.02		0.03		0.04
			(0.08)		(0.07)		(0.07)
25-29			0.06		0.09		-0.02
			(0.09)		(0.09)		(0.09)

Weighted regressions. Standard errors clustered at the classroom level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. (*Continued on next page*)

Table 7: Second stage of Class FE Analysis (Maths) (continued)

	Dependent variable: Class fixed effect						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
> 30			0.13		0.14		0.11
			(0.09)		(0.09)		(0.09)
Received training			-0.06**		-0.08***		-0.05*
			(0.03)		(0.03)		(0.03)
Teacher characteristics:							
Female				-0.02	-0.00	-0.01	-0.01
				(0.04)	(0.04)	(0.04)	(0.03)
Experience (years):							
5 - 9				-0.13*	-0.14**	-0.12*	-0.14**
				(0.07)	(0.06)	(0.07)	(0.06)
10 - 14				-0.02	-0.06	-0.04	-0.04
				(0.09)	(0.08)	(0.09)	(0.08)
15 - 19				0.06	0.07	0.05	0.07
				(0.08)	(0.06)	(0.08)	(0.06)
20 - 24				0.06	0.04	0.05	0.04
				(0.07)	(0.06)	(0.08)	(0.06)
25 - 29				0.06	0.09	0.04	0.05
				(0.07)	(0.06)	(0.07)	(0.05)
30 or more				0.07	0.13**	0.06	0.09*
				(0.06)	(0.05)	(0.06)	(0.05)
5-years degree or more				-0.04	0.01	-0.04	-0.00
				(0.04)	(0.04)	(0.04)	(0.04)
Taught subjects:							
Reading and Maths				-0.13	-0.09*	-0.12	-0.01
				(0.08)	(0.05)	(0.08)	(0.06)
Reading				-0.11	0.00	-0.13	0.09
				(0.09)	(0.08)	(0.10)	(0.08)
Training:							
Attending courses				0.00	0.02	0.01	0.02
				(0.04)	(0.03)	(0.04)	(0.03)
Working teams at school				-0.04	-0.01	-0.03	0.02
				(0.04)	(0.03)	(0.04)	(0.03)
Congresses and teaching projects				-0.01	-0.05*	0.01	-0.08***
				(0.04)	(0.03)	(0.04)	(0.03)
Type of warnings to students:							
Letter to the family				-0.06	-0.09***	-0.07*	-0.06*

Weighted regressions. Standard errors clustered at the classroom level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. (*Continued on next page*)

Table 7: Second stage of Class FE Analysis (Maths) (continued)

	Dependent variable: Class fixed effect						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				(0.04)	(0.03)	(0.04)	(0.03)
Temporary class suspension				-0.09*	-0.11***	-0.09	-0.07*
				(0.05)	(0.04)	(0.05)	(0.04)
Person asking for a meeting:							
Parents				-0.02	0.09**	-0.01	0.06
				(0.05)	(0.04)	(0.06)	(0.04)
Teacher				0.03	-0.04	0.03	-0.01
				(0.05)	(0.03)	(0.05)	(0.03)
# meetings with parents				-0.06*	-0.02	-0.05	0.00
				(0.03)	(0.02)	(0.03)	(0.02)
Teacher at 3^{rd} and 4^{th} grades				0.08*	0.05	0.08*	0.04
				(0.04)	(0.03)	(0.04)	(0.03)
Class characteristics:							
Class size				0.00	0.02***	0.00	0.01**
				(0.00)	(0.00)	(0.01)	(0.00)
Mother's education:							
Compulsory						0.15	0.32*
						(0.22)	(0.18)
High School						0.34	0.43**
						(0.27)	(0.20)
Vocational training						0.24	0.75***
						(0.28)	(0.20)
University						0.23	0.75***
						(0.28)	(0.22)
Father's education:							
Compulsory						0.11	0.02
						(0.27)	(0.19)
High School						0.03	-0.10
						(0.27)	(0.20)
Vocational training						0.34	0.25
						(0.26)	(0.19)
University						-0.07	-0.07
						(0.28)	(0.21)
Percent non-Spanish							
10-20%						-0.02	0.05
						(0.05)	(0.04)

Weighted regressions. Standard errors clustered at the classroom level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. (*Continued on next page*)

Table 7: Second stage of Class FE Analysis (Maths) (continued)

	Dependent variable: Class fixed effect						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
More 20%						-0.02	0.15***
						(0.06)	(0.05)
Mother's labor status:							
Employee						-0.01	0.00
						(0.17)	(0.15)
Unemployed						0.33	-0.13
						(0.25)	(0.21)
Inactive						0.38	0.14
						(0.25)	(0.16)
Father's labor status:							
Employee						-0.03	-0.15
						(0.15)	(0.11)
Unemployed						-0.06	-0.32*
						(0.25)	(0.19)
Inactive						0.51	-0.52
						(0.43)	(0.37)
Percent single-parent						0.05	-0.32*
						(0.23)	(0.20)
Percent siblings						-0.13	-0.25**
						(0.15)	(0.12)
Percent female						0.07	0.08
						(0.14)	(0.11)
Percent repeater						-0.16	-0.18
						(0.19)	(0.13)
Constant	0.01	0.61***	-0.84**	0.87***	-0.61*	0.62	-0.47
	(0.11)	(0.18)	(0.40)	(0.25)	(0.33)	(0.43)	(0.45)
School characteristics	No	No	Yes	No	Yes	No	Yes
Principal characteristics	No	No	Yes	No	Yes	No	Yes
School fixed effects	No	Yes	No	No	No	Yes	No
Teacher characteristics	No	No	No	Yes	Yes	Yes	Yes
Class characteristics	No	No	No	Yes	No	Yes	Yes
Classrooms	748	748	732	748	732	748	732
R^2	0.00	0.75	0.09	0.77	0.25	0.78	0.37

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

Table 8: Second stage of Class FE Analysis (Reading)

	Dependent variable: Class fixed effect						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Traditional Teaching Practices	0.01	-0.00	0.01	-0.01	-0.02	-0.00	0.00
	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)
Modern Teaching Practices	-0.06	0.05	-0.03	0.05	-0.03	0.05	-0.03
	(0.04)	(0.05)	(0.04)	(0.05)	(0.04)	(0.05)	(0.03)
School characteristics:							
Public school			-0.03		-0.02		0.02
			(0.09)		(0.09)		(0.09)
Size of municipality:							
2000-10000			0.29**		0.31*		0.33**
			(0.13)		(0.17)		(0.15)
10000-50000			0.23*		0.23		0.27*
			(0.13)		(0.17)		(0.15)
50000-500000			0.24*		0.25		0.29**
			(0.13)		(0.17)		(0.15)
$>$ 500000			0.38***		0.36**		0.36**
			(0.14)		(0.17)		(0.15)
Education levels at school:							
Preschool			-0.05		-0.03		-0.00
			(0.09)		(0.09)		(0.08)
Primary			0.09		-0.06		-0.20
			(0.13)		(0.18)		(0.17)
Compulsory			0.01		-0.08		-0.10
			(0.10)		(0.10)		(0.10)
Academic high school			-0.01		-0.01		-0.01
			(0.06)		(0.06)		(0.06)
Vocational high school			-0.03		0.00		0.01
			(0.13)		(0.13)		(0.12)
Vocational college			0.10		0.04		0.01
			(0.15)		(0.17)		(0.15)
Teaching hours per week			0.01***		0.01***		0.01**
			(0.00)		(0.00)		(0.00)
Rule to group students:							
Alfabethic			0.10***		0.08**		0.07**
			(0.04)		(0.03)		(0.03)
Gender equilibrium			0.10**		0.06*		0.04

Weighted regressions. Standard errors clustered at the classroom level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. (*Continued on next page*)

Table 8: Second stage of Class FE Analysis (Reading) (continued)

	Dependent variable: Class fixed effect						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			(0.04)		(0.04)		(0.04)
Language			-0.03		-0.02		-0.05
			(0.04)		(0.04)		(0.04)
Academic outcomes			-0.03		-0.04		-0.03
			(0.05)		(0.04)		(0.04)
Searching homogeneity			0.03		0.02		0.03
			(0.04)		(0.04)		(0.04)
Searching heterogeneity			0.04		0.03		0.02
			(0.04)		(0.03)		(0.03)
Others			0.02		-0.01		-0.03
			(0.04)		(0.03)		(0.03)
Principals in last 10 years:							
2			0.06		0.05		0.02
			(0.05)		(0.05)		(0.04)
3			0.01		0.00		-0.03
			(0.05)		(0.05)		(0.05)
> 3			0.04		0.04		0.00
			(0.06)		(0.06)		(0.06)
Principal characteristics:							
Female			0.06*		0.07**		0.04
			(0.03)		(0.03)		(0.03)
Non-university			0.08		0.04		0.03
			(0.06)		(0.06)		(0.06)
5-years degree			0.07**		0.06*		0.05
			(0.03)		(0.03)		(0.03)
Experience (years):							
5-9			0.03		0.06		0.04
			(0.04)		(0.04)		(0.04)
10-14			0.04		0.03		0.02
			(0.05)		(0.05)		(0.05)
15-19			-0.06		-0.05		-0.04
			(0.06)		(0.06)		(0.05)
20-24			0.10		0.10		0.12*
			(0.06)		(0.06)		(0.06)
25-29			0.06		0.08		-0.03
			(0.09)		(0.10)		(0.09)

Weighted regressions. Standard errors clustered at the classroom level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. (*Continued on next page*)

Table 8: Second stage of Class FE Analysis (Reading) (continued)

	Dependent variable: Class fixed effect						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
> 30			0.23***		0.22**		0.21**
			(0.08)		(0.10)		(0.09)
Received training			-0.06**		-0.08***		-0.05*
			(0.03)		(0.03)		(0.03)
Teacher characteristics:							
Female				0.02	-0.01	0.01	-0.02
				(0.04)	(0.04)	(0.04)	(0.03)
Experience (years):							
5 - 9				0.05	-0.01	0.05	-0.01
				(0.07)	(0.07)	(0.07)	(0.06)
10 - 14				0.14*	0.04	0.12	0.07
				(0.08)	(0.07)	(0.08)	(0.06)
15 - 19				0.05	0.08	0.04	0.08
				(0.08)	(0.06)	(0.08)	(0.06)
20 - 24				-0.01	0.05	-0.00	0.06
				(0.07)	(0.06)	(0.07)	(0.06)
25 - 29				0.09	0.12**	0.09	0.09*
				(0.07)	(0.06)	(0.07)	(0.05)
30 or more				0.10*	0.15***	0.11*	0.11**
				(0.06)	(0.05)	(0.06)	(0.05)
5-years degree or more				-0.09**	-0.03	-0.09**	-0.05
				(0.04)	(0.04)	(0.04)	(0.04)
Taught subjects:							
Reading and Maths				-0.10	-0.07	-0.12	0.00
				(0.09)	(0.06)	(0.10)	(0.06)
Reading				-0.07	0.02	-0.10	0.08
				(0.10)	(0.08)	(0.10)	(0.07)
Training:							
Attending courses				-0.04	-0.02	-0.03	-0.02
				(0.04)	(0.03)	(0.04)	(0.03)
Working teams at school				-0.04	-0.02	-0.02	0.02
				(0.04)	(0.03)	(0.04)	(0.03)
Congresses and teaching projects				0.03	0.03	0.03	0.00
				(0.03)	(0.03)	(0.04)	(0.03)
Type of warnings to students:							
Letter to the family				0.00	-0.04	-0.00	-0.01

Weighted regressions. Standard errors clustered at the classroom level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. (*Continued on next page*)

Table 8: Second stage of Class FE Analysis (Reading) (continued)

	Dependent variable: Class fixed effect						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				(0.04)	(0.03)	(0.04)	(0.03)
Temporary class suspension				-0.02	-0.06	-0.01	-0.03
				(0.05)	(0.04)	(0.05)	(0.04)
Person asking for a meeting:							
Parents				-0.05	0.04	-0.05	0.01
				(0.05)	(0.04)	(0.05)	(0.03)
Teacher				0.04	-0.04	0.05	-0.02
				(0.05)	(0.03)	(0.05)	(0.03)
# meetings with parents				-0.05	-0.01	-0.04	0.01
				(0.03)	(0.02)	(0.03)	(0.02)
Teacher at 3^{rd} and 4^{th} grades				0.04	-0.02	0.04	-0.03
				(0.04)	(0.04)	(0.04)	(0.03)
Class characteristics:							
Class size				0.01**	0.03***	0.01***	0.02***
				(0.00)	(0.00)	(0.00)	(0.00)
Mother's education:							
Compulsory						0.23	0.51***
						(0.22)	(0.19)
High School						0.49*	0.54**
						(0.26)	(0.21)
Vocational training						0.45*	0.79***
						(0.25)	(0.21)
University						0.21	0.80***
						(0.28)	(0.23)
Father's education:							
Compulsory						-0.11	0.07
						(0.24)	(0.18)
High School						-0.38	-0.22
						(0.26)	(0.20)
Vocational training						0.05	0.08
						(0.25)	(0.19)
University						-0.39	-0.11
						(0.26)	(0.21)
Percent non-Spanish							
10-20%						0.02	0.08*
						(0.05)	(0.04)

Weighted regressions. Standard errors clustered at the classroom level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. (*Continued on next page*)

Table 8: Second stage of Class FE Analysis (Reading) (continued)

	Dependent variable: Class fixed effect						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
More 20%						0.05	0.21***
						(0.07)	(0.05)
Mother's labor status:							
Employee						-0.03	0.05
						(0.16)	(0.17)
Unemployed						-0.12	-0.13
						(0.23)	(0.22)
Inactive						-0.07	0.19
						(0.20)	(0.18)
Father's labor status:							
Employee						0.02	-0.04
						(0.15)	(0.12)
Unemployed						-0.08	-0.48**
						(0.27)	(0.21)
Inactive						-0.37	-0.82**
						(0.42)	(0.39)
Percent single-parent						-0.12	-0.23
						(0.25)	(0.19)
Percent siblings						-0.26*	-0.27**
						(0.15)	(0.11)
Percent female						-0.05	-0.02
						(0.14)	(0.10)
Percent repeater						-0.03	-0.03
						(0.15)	(0.14)
Constant	-0.00	0.34**	-0.80***	0.34	-0.79**	0.56	-1.00**
	(0.12)	(0.16)	(0.26)	(0.25)	(0.32)	(0.46)	(0.43)
School characteristics	No	No	Yes	No	Yes	No	Yes
Principal characteristics	No	No	Yes	No	Yes	No	Yes
School fixed effects	No	Yes	No	No	No	Yes	No
Teacher characteristics	No	No	No	Yes	Yes	Yes	Yes
Class characteristics	No	No	No	Yes	No	Yes	Yes
Classrooms	748	748	732	748	732	748	732
R^2	0.00	0.76	0.10	0.78	0.23	0.80	0.36

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

Table 9: Rules used by school to assign students to classrooms (%)

Alfabethic or other random rule	44.19
Language abilities	14.21
Academic ability	11.11
Heterogeneous classes	56.85
Well mixed classes (gender, origin)	77.26
Other rules	24.55
Number of schools	387

Percentage of principals who chose the respective rule. Multiple answers were possible

Table 10: Balancing tests

	Mother's education			
	Compulsory	High School	Vocational training	University
Traditional Teaching Practices	-0.00	-0.00	-0.03***	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)
Modern Teaching Practices	-0.01	-0.00	0.02*	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Observations	12785	12785	12785	12785
R^2	0.11	0.05	0.06	0.21
F-test	0.14	0.32	6.82	4.48
p-value	0.87	0.73	0.00	0.01
	Father's education			
	Compulsory	High School	Vocational training	University
Traditional Teaching Practices	-0.00	0.01	0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Modern Teaching Practices	-0.00	-0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Observations	12785	12785	12785	12785
R^2	0.10	0.04	0.06	0.21
F-test	0.09	1.35	0.90	0.58
p-value	0.92	0.26	0.41	0.56
	Single-parent household	Siblings	Female	Repeater
Traditional Teaching Practices	-0.00	0.01	0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Modern Teaching Practices	-0.02**	0.00	0.02	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Observations	12785	12785	12785	12785
R^2	0.04	0.06	0.05	0.05
F-test	3.28	0.77	0.80	0.84
p-value	0.04	0.46	0.45	0.43
		Mother's labor status		
	Non-Spanish origin	Employee	Unemployed	Inactive
Traditional Teaching Practices	-0.01*	0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)
Modern Teaching Practices	-0.00	-0.00	-0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Observations	12785	12785	12785	12785
R^2	0.14	0.10	0.06	0.09
F-test	1.86	0.03	1.37	0.76
p-value	0.16	0.97	0.25	0.47
	Father's labor status			
	Employee	Unemployed	Inactive	
Traditional Teaching Practices	0.01	-0.00	-0.00	
	(0.01)	(0.01)	(0.00)	
Modern Teaching Practices	-0.03**	-0.02***	-0.00	
	(0.01)	(0.01)	(0.00)	
Observations	12785	12785	12785	
R^2	0.05	0.08	0.04	
F-test	3.71	4.41	0.41	
p-value	0.02	0.01	0.66	

Each column in each panel represents a separate regression. Standard errors clustered at the classroom level in parentheses. School fixed effect included in all regressions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Teaching practices and class size

	Traditional teaching practices		Modern teaching practices	
	OLS	School FE	OLS	School FE
Class size	0.02	0.00	-0.01	0.03
	(0.02)	(0.02)	(0.02)	(0.02)
Class size^2	-0.00	-0.00	0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Observations	12785	12785	12785	12785
R^2	0.00	0.68	0.00	0.70
F-test	2.19	0.00	0.45	1.26
p-value	0.11	1.00	0.64	0.29

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

Table 12: Effect of class-average characteristics

	Dependent variable: Teacher characteristics							
		Years of experience						
	Female	5 - 9	10 - 14	15 - 19	20 - 24	25 - 29	30 or more	5-year degree
Mother's education:								
Compulsory	0.16	0.15	0.19	-0.01	-0.29**	-0.28*	0.35	0.17
	(0.18)	(0.15)	(0.13)	(0.15)	(0.13)	(0.17)	(0.25)	(0.15)
High School	0.29	-0.08	0.49***	-0.10	-0.04	-0.37*	0.16	0.30
	(0.24)	(0.18)	(0.13)	(0.16)	(0.19)	(0.21)	(0.30)	(0.21)
Vocational training	0.05	-0.01	0.36**	-0.05	-0.23	-0.22	0.30	0.14
	(0.23)	(0.17)	(0.14)	(0.15)	(0.20)	(0.22)	(0.31)	(0.21)
University	0.15	-0.03	0.38***	0.11	-0.42**	-0.38*	0.30	0.10
	(0.24)	(0.17)	(0.14)	(0.16)	(0.19)	(0.21)	(0.31)	(0.23)
Father's education:								
Compulsory	-0.25	-0.14	-0.03	-0.07	0.59***	0.24	-0.30	-0.21
	(0.23)	(0.15)	(0.14)	(0.17)	(0.18)	(0.18)	(0.29)	(0.18)
High School	-0.53**	-0.07	-0.45***	0.07	0.65***	0.52***	-0.29	-0.34
	(0.23)	(0.16)	(0.13)	(0.14)	(0.20)	(0.20)	(0.28)	(0.22)
Vocational training	-0.32	0.02	-0.26*	0.11	0.47**	0.55**	-0.73***	-0.53**
	(0.24)	(0.15)	(0.14)	(0.16)	(0.19)	(0.22)	(0.28)	(0.21)
University	-0.19	0.17	-0.15	-0.22	0.70***	0.35	-0.69**	-0.14
	(0.24)	(0.16)	(0.14)	(0.16)	(0.19)	(0.22)	(0.29)	(0.23)
Percent non-Spanish								
10-20%	0.00	0.04	-0.00	-0.07**	-0.05*	-0.02	0.07	0.13***
	(0.05)	(0.03)	(0.02)	(0.03)	(0.03)	(0.05)	(0.06)	(0.04)
More 20%	-0.08	0.08	0.09**	-0.02	-0.08	-0.01	-0.11	0.24***
	(0.07)	(0.06)	(0.04)	(0.04)	(0.05)	(0.06)	(0.09)	(0.07)
Mother's labor status:								
Employee	0.49***	-0.02	-0.01	0.04	-0.10	0.19	-0.34*	0.06
	(0.17)	(0.12)	(0.12)	(0.13)	(0.12)	(0.16)	(0.19)	(0.18)
Unemployed	0.14	0.03	0.16	-0.26	-0.08	-0.13	0.29	0.23
	(0.21)	(0.19)	(0.16)	(0.17)	(0.16)	(0.19)	(0.27)	(0.24)
Inactive	0.19	0.02	0.01	-0.04	-0.22	0.39**	-0.38	0.05
	(0.19)	(0.13)	(0.12)	(0.15)	(0.14)	(0.19)	(0.25)	(0.19)
Father's labor status:								
Employee	0.10	0.07	-0.03	0.14	-0.29***	0.03	0.08	-0.33**
	(0.15)	(0.09)	(0.08)	(0.10)	(0.10)	(0.13)	(0.16)	(0.14)
Unemployed	-0.44*	-0.00	0.13	0.41**	0.12	0.08	-0.79***	-0.83***
	(0.24)	(0.19)	(0.16)	(0.17)	(0.19)	(0.22)	(0.28)	(0.24)
Inactive	-0.83*	-0.48*	0.37	0.07	-0.31	0.75*	-0.48	0.08
	(0.45)	(0.26)	(0.27)	(0.24)	(0.35)	(0.45)	(0.60)	(0.37)
Percent single-parent	0.19	-0.05	0.13	0.05	-0.27	-0.07	0.39	-0.03
	(0.22)	(0.15)	(0.14)	(0.14)	(0.20)	(0.21)	(0.29)	(0.24)
Percent siblings	-0.07	0.17	0.01	-0.09	0.01	-0.13	0.21	-0.23
	(0.14)	(0.11)	(0.08)	(0.08)	(0.12)	(0.13)	(0.16)	(0.16)
Percent female	-0.39***	0.12	0.00	0.01	-0.02	-0.01	-0.07	-0.04
	(0.14)	(0.09)	(0.08)	(0.09)	(0.10)	(0.13)	(0.16)	(0.13)
Percent repeater	0.06	0.20*	-0.01	-0.18	-0.06	-0.03	0.09	0.26*
	(0.19)	(0.10)	(0.13)	(0.12)	(0.13)	(0.16)	(0.21)	(0.15)
Observations	12785	12785	12785	12785	12785	12785	12785	12785
R^2	0.59	0.57	0.57	0.62	0.53	0.55	0.56	0.54
F-test	2.44	1.55	1.97	1.47	2.08	1.42	2.21	2.68
p-value	0.00	0.06	0.01	0.09	0.00	0.11	0.00	0.00

Each column represents a separate regression. Standard errors clustered at the classroom level in parentheses. School fixed effect included in all regressions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13: Effect of class-average characteristics

	Dependent variable: Teacher characteristics							
	Teaching Practices		Taught subjects		Training			
	Traditional	Modern	Read-Maths	Read	Courses	Work teams	Congress	3^{rd}-4^{th}
Mother's education:								
Compulsory	-0.19	-0.24	-0.22	0.06	-0.11	0.00	0.35	0.20
	(0.20)	(0.17)	(0.16)	(0.14)	(0.21)	(0.21)	(0.23)	(0.18)
High School	-0.35	-0.27	-0.14	0.07	0.15	0.38	0.53**	0.11
	(0.25)	(0.19)	(0.16)	(0.15)	(0.24)	(0.26)	(0.26)	(0.22)
Vocational training	-0.69***	-0.03	-0.21	0.06	-0.22	0.42*	0.27	0.03
	(0.26)	(0.20)	(0.18)	(0.15)	(0.25)	(0.24)	(0.27)	(0.24)
University	0.16	-0.37*	-0.17	-0.00	-0.29	-0.06	-0.07	-0.34
	(0.26)	(0.20)	(0.18)	(0.16)	(0.26)	(0.25)	(0.27)	(0.25)
Father's education:								
Compulsory	0.13	0.08	0.54***	-0.32**	0.15	-0.36	-0.77***	0.11
	(0.22)	(0.20)	(0.14)	(0.13)	(0.26)	(0.23)	(0.26)	(0.22)
High School	0.36	-0.13	0.07	-0.00	0.28	-0.03	-0.06	0.32
	(0.23)	(0.19)	(0.13)	(0.14)	(0.26)	(0.24)	(0.27)	(0.20)
Vocational training	0.31	0.11	0.22	-0.16	0.08	-0.42*	-0.76***	-0.11
	(0.21)	(0.19)	(0.15)	(0.14)	(0.24)	(0.22)	(0.26)	(0.23)
University	-0.12	0.15	0.11	-0.15	0.49*	0.11	-0.21	0.25
	(0.22)	(0.20)	(0.13)	(0.13)	(0.26)	(0.24)	(0.25)	(0.22)
Percent non-Spanish								
10-20%	-0.01	-0.03	-0.01	-0.00	-0.01	-0.15***	0.10*	0.00
	(0.05)	(0.04)	(0.02)	(0.01)	(0.05)	(0.05)	(0.05)	(0.04)
More 20%	-0.04	0.11*	-0.01	0.02	-0.03	-0.13*	0.14**	-0.06
	(0.07)	(0.06)	(0.05)	(0.04)	(0.08)	(0.07)	(0.07)	(0.07)
Mother's labor status:								
Employee	0.18	0.18	-0.17*	0.09	-0.04	-0.13	-0.12	-0.08
	(0.17)	(0.16)	(0.09)	(0.08)	(0.20)	(0.15)	(0.18)	(0.16)
Unemployed	0.30	-0.11	-0.50***	0.33***	-0.16	-0.53**	-0.13	-0.25
	(0.23)	(0.18)	(0.16)	(0.13)	(0.25)	(0.22)	(0.25)	(0.21)
Inactive	0.30	0.12	-0.12	0.06	0.04	-0.33*	-0.49**	-0.09
	(0.20)	(0.18)	(0.12)	(0.10)	(0.23)	(0.19)	(0.21)	(0.19)
Father's labor status:								
Employee	0.04	-0.51***	0.21**	-0.04	-0.01	-0.04	-0.28*	0.11
	(0.13)	(0.11)	(0.10)	(0.07)	(0.15)	(0.14)	(0.15)	(0.12)
Unemployed	-0.24	-0.82***	0.01	-0.02	0.47*	0.16	-0.14	-0.07
	(0.21)	(0.19)	(0.14)	(0.09)	(0.27)	(0.26)	(0.26)	(0.25)
Inactive	-0.23	-0.54*	0.49**	-0.55***	-0.22	0.46	-0.74	0.07
	(0.41)	(0.32)	(0.25)	(0.20)	(0.54)	(0.41)	(0.49)	(0.42)
Percent single-parent	-0.09	-0.48**	0.12	-0.01	-0.16	0.18	-0.15	0.22
	(0.21)	(0.20)	(0.14)	(0.11)	(0.25)	(0.25)	(0.23)	(0.23)
Percent siblings	0.16	-0.05	0.10	-0.17**	0.17	0.06	0.01	0.25
	(0.16)	(0.12)	(0.10)	(0.07)	(0.16)	(0.15)	(0.17)	(0.15)
Percent female	-0.01	0.08	-0.10	-0.04	0.24*	0.45***	0.53***	0.21*
	(0.13)	(0.13)	(0.09)	(0.09)	(0.14)	(0.13)	(0.15)	(0.12)
Percent repeater	0.24	-0.01	0.21**	-0.07	0.10	0.11	-0.14	0.02
	(0.16)	(0.14)	(0.10)	(0.08)	(0.20)	(0.19)	(0.19)	(0.18)
Observations	12785	12785	12785	12785	12785	12785	12785	12675
R^2	0.70	0.72	0.72	0.57	0.58	0.69	0.65	0.64
F-test	1.96	3.25	2.34	1.41	0.89	2.63	3.67	1.28
p-value	0.01	0.00	0.00	0.11	0.60	0.00	0.00	0.18

Each column represents a separate regression. Standard errors clustered at the classroom level in parentheses. School fixed effect included in all regressions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 14: Estimation results of deviation to school average

	Tutor			Students		
	(1)	(2)	(3)	(4)	(5)	(6)
Traditional Teaching Practices	0.01	-0.00	-0.00	0.07***	0.07***	0.06***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
Modern Teaching Practices	0.05	0.04	0.05	-0.25***	-0.24***	-0.20***
	(0.04)	(0.04)	(0.04)	(0.01)	(0.01)	(0.01)
Maths dummy	-0.00	0.00	0.00	0.03***	0.01***	0.01**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Class size		0.00	-0.00		-0.00***	-0.00***
		(0.00)	(0.00)		(0.00)	(0.00)
Female		0.02	0.01		0.03	0.02
		(0.03)	(0.03)		(0.03)	(0.03)
Experience (years):						
5 - 9		-0.01	-0.00		-0.01	0.00
		(0.05)	(0.05)		(0.05)	(0.05)
10 - 14		0.06	0.07		0.05	0.06
		(0.07)	(0.07)		(0.07)	(0.07)
15 - 19		0.08	0.08		0.04	0.05
		(0.06)	(0.06)		(0.06)	(0.06)
20 - 24		0.01	0.01		-0.02	-0.02
		(0.06)	(0.05)		(0.05)	(0.05)
25 - 29		0.07	0.07		0.03	0.04
		(0.06)	(0.05)		(0.05)	(0.05)
30 or more		0.07	0.08*		0.05	0.06
		(0.05)	(0.05)		(0.05)	(0.04)
5-years degree or more		-0.09***	-0.07**		-0.08**	-0.07**
		(0.03)	(0.03)		(0.03)	(0.03)
Taught subjects:						
Reading and Maths		-0.11*	-0.06		-0.05	-0.02
		(0.06)	(0.06)		(0.06)	(0.06)
Reading		-0.13*	-0.10		-0.10	-0.07
		(0.07)	(0.07)		(0.07)	(0.07)
Training:						
Attending courses		0.01	0.00		0.03	0.02
		(0.03)	(0.03)		(0.03)	(0.03)
Working teams at school		-0.03	-0.04		-0.02	-0.03
		(0.03)	(0.03)		(0.04)	(0.04)

Dependent variable: Deviation to school-average test scores. Standard errors clustered at the school level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. (*Continued on next page*)

Table 14: (continued)

	Tutor			Students		
	(1)	(2)	(3)	(4)	(5)	(6)
Congresses and teaching projects		0.01	0.01		0.02	0.02
		(0.03)	(0.03)		(0.03)	(0.03)
Type of warnings to students:						
Letter to the family		-0.03	-0.03		-0.04	-0.03
		(0.03)	(0.03)		(0.03)	(0.03)
Temporary class suspension		-0.04	-0.04		-0.05	-0.05
		(0.04)	(0.04)		(0.05)	(0.05)
Person asking for a meeting:						
Parents		-0.01	-0.00		-0.01	-0.01
		(0.04)	(0.04)		(0.04)	(0.04)
Teacher		0.00	0.03		0.01	0.02
		(0.04)	(0.04)		(0.04)	(0.04)
Number of meetings with students' parents		-0.03	-0.03		-0.02	-0.02
		(0.03)	(0.03)		(0.03)	(0.03)
Teacher at 3^{rd} and 4^{th} grades		0.04	0.04		0.06*	0.05*
		(0.03)	(0.03)		(0.03)	(0.03)
Contextual characteristics	No	No	Yes	No	No	Yes
Observations	25570	25094	25094	23270	22834	22834
R^2	0.00	0.00	0.06	0.03	0.03	0.08

Dependent variable: Deviation to school-average test scores. Standard errors clustered at the school level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 15: Estimation results of deviation to school average

	Tutor			Students		
	(1)	(2)	(3)	(4)	(5)	(6)
Traditional Teaching Materials	0.03	0.03	0.03	0.06***	0.06***	0.05***
	(0.03)	(0.03)	(0.03)	(0.01)	(0.01)	(0.01)
Modern Teaching Materials	0.02	0.02	0.02	-0.21***	-0.21***	-0.18***
	(0.03)	(0.03)	(0.04)	(0.01)	(0.01)	(0.01)
Maths dummy	-0.00	-0.00	-0.00	0.02***	0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Class size		-0.00	-0.00		-0.00***	-0.00***
		(0.00)	(0.00)		(0.00)	(0.00)
Female		0.02	0.01		0.02	0.01
		(0.03)	(0.03)		(0.02)	(0.03)
Experience (years):						
5 - 9		-0.01	-0.00		0.00	0.00
		(0.05)	(0.05)		(0.05)	(0.05)
10 - 14		0.06	0.06		0.07	0.08
		(0.07)	(0.07)		(0.07)	(0.07)
15 - 19		0.07	0.07		0.09	0.09*
		(0.06)	(0.06)		(0.06)	(0.05)
20 - 24		0.00	0.01		-0.01	-0.01
		(0.06)	(0.05)		(0.05)	(0.05)
25 - 29		0.07	0.07		0.06	0.07
		(0.06)	(0.06)		(0.05)	(0.05)
30 or more		0.07	0.07		0.06	0.07
		(0.05)	(0.05)		(0.04)	(0.04)
5-years degree or more		-0.09***	-0.07**		-0.10***	-0.08***
		(0.03)	(0.03)		(0.03)	(0.03)
Taught subjects:						
Reading and Maths		-0.11*	-0.06		-0.09	-0.06
		(0.06)	(0.06)		(0.06)	(0.06)
Reading		-0.13**	-0.10		-0.12*	-0.10
		(0.07)	(0.07)		(0.07)	(0.07)
Training:						
Attending courses		0.01	0.00		0.02	0.01
		(0.03)	(0.03)		(0.03)	(0.03)
Working teams at school		-0.03	-0.04		-0.03	-0.03
		(0.03)	(0.03)		(0.03)	(0.03)

Dependent variable: Deviation to school-average test scores. Standard errors clustered at the school level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. (*Continued on next page*)

Table 15: (continued)

	Tutor			Students		
	(1)	(2)	(3)	(4)	(5)	(6)
Congresses and teaching projects		0.01	0.02		0.02	0.02
		(0.03)	(0.03)		(0.03)	(0.03)
Type of warnings to students:						
Letter to the family		-0.03	-0.03		-0.03	-0.03
		(0.03)	(0.03)		(0.03)	(0.03)
Temporary class suspension		-0.04	-0.05		-0.04	-0.04
		(0.05)	(0.04)		(0.05)	(0.04)
Person asking for a meeting:						
Parents		-0.01	-0.00		-0.01	-0.00
		(0.04)	(0.04)		(0.04)	(0.04)
Teacher		0.01	0.03		0.01	0.04
		(0.04)	(0.04)		(0.04)	(0.04)
Number of meetings with students' parents		-0.04	-0.04		-0.02	-0.02
		(0.03)	(0.03)		(0.03)	(0.03)
Teacher at 3^{rd} and 4^{th} grades		0.05	0.05		0.05*	0.05
		(0.03)	(0.03)		(0.03)	(0.03)
Contextual characteristics	No	No	Yes	No	No	Yes
Observations	25132	24728	24728	23768	23336	23336
R^2	0.00	0.00	0.06	0.03	0.03	0.08

Dependent variable: Deviation to school-average test scores. Standard errors clustered at the school level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 16: Estimation results of deviation to school average (Maths)

	Tutor	Students	Tutor	Students
	(1)	(2)	(3)	(4)
Traditional Teaching Practices	0.02	0.06***		
	(0.04)	(0.02)		
Modern Teaching Practices	0.02	-0.18***		
	(0.05)	(0.01)		
Traditional Teaching Materials			0.03	0.04***
			(0.04)	(0.01)
Modern Teaching Materials			-0.01	-0.17***
			(0.04)	(0.01)
Class size	-0.00	-0.00***	-0.00	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
Female	0.00	0.02	0.00	0.00
	(0.03)	(0.03)	(0.03)	(0.03)
Experience (years):				
5 - 9	-0.07	-0.07	-0.07	-0.07
	(0.06)	(0.06)	(0.06)	(0.06)
10 - 14	0.03	0.04	0.03	0.04
	(0.08)	(0.08)	(0.08)	(0.08)
15 - 19	0.10	0.06	0.10	0.11
	(0.07)	(0.07)	(0.07)	(0.07)
20 - 24	0.05	0.02	0.05	0.03
	(0.06)	(0.06)	(0.06)	(0.06)
25 - 29	0.08	0.05	0.08	0.08
	(0.06)	(0.06)	(0.06)	(0.06)
30 or more	0.09	0.07	0.08	0.08
	(0.06)	(0.05)	(0.06)	(0.05)
5-years degree or more	-0.05	-0.06*	-0.06*	-0.06*
	(0.03)	(0.03)	(0.03)	(0.03)
Taught subjects:				
Reading and Maths	-0.09	-0.06	-0.09	-0.08
	(0.07)	(0.07)	(0.07)	(0.07)
Reading	-0.13	-0.14	-0.14*	-0.14*
	(0.08)	(0.09)	(0.08)	(0.08)
Training:				
Attending courses	0.03	0.05	0.03	0.04

Dependent variable: Deviation to school-average test scores. Standard errors clustered at the school level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
(*Continued on next page*)

Table 16: (continued)

	Tutor	Students	Tutor	Students
	(1)	(2)	(3)	(4)
	(0.04)	(0.04)	(0.04)	(0.04)
Working teams at school	-0.04	-0.04	-0.04	-0.04
	(0.04)	(0.04)	(0.04)	(0.04)
Congresses and teaching projects	0.00	0.01	0.01	0.01
	(0.03)	(0.03)	(0.03)	(0.03)
Type of warnings to students:				
Letter to the family	-0.06	-0.06*	-0.06	-0.06*
	(0.03)	(0.04)	(0.03)	(0.03)
Temporary class suspension	-0.06	-0.07	-0.07	-0.06
	(0.05)	(0.06)	(0.05)	(0.05)
Person asking for a meeting:				
Parents	0.02	-0.00	0.01	0.01
	(0.05)	(0.05)	(0.05)	(0.05)
Teacher	0.02	0.02	0.03	0.03
	(0.05)	(0.05)	(0.05)	(0.05)
Number of meetings with students' parents	-0.04	-0.02	-0.04	-0.02
	(0.03)	(0.03)	(0.03)	(0.03)
Teacher at 3^{rd} and 4^{th} grades	0.06	0.06*	0.06	0.06
	(0.04)	(0.04)	(0.04)	(0.04)
Contextual characteristics	Yes	Yes	Yes	Yes
Observations	12547	11417	12364	11668
R^2	0.06	0.08	0.06	0.08

Dependent variable: Deviation to school-average test scores. Standard errors clustered at the school level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 17: Estimation results of deviation to school average (Reading)

	Tutor	Students	Tutor	Students
	(1)	(2)	(3)	(4)
Traditional Teaching Practices	-0.02	0.07***		
	(0.04)	(0.02)		
Modern Teaching Practices	0.08*	-0.21***		
	(0.04)	(0.01)		
Traditional Teaching Materials			0.02	0.06***
			(0.03)	(0.01)
Modern Teaching Materials			0.04	-0.19***
			(0.04)	(0.01)
Class size	-0.00	-0.00***	-0.00	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
Female	0.02	0.02	0.02	0.01
	(0.03)	(0.03)	(0.03)	(0.03)
Experience (years):				
5 - 9	0.07	0.07	0.06	0.07
	(0.06)	(0.06)	(0.06)	(0.06)
10 - 14	0.11	0.09	0.10	0.11
	(0.07)	(0.08)	(0.08)	(0.07)
15 - 19	0.05	0.04	0.05	0.08
	(0.07)	(0.07)	(0.07)	(0.07)
20 - 24	-0.04	-0.06	-0.04	-0.06
	(0.07)	(0.07)	(0.07)	(0.07)
25 - 29	0.06	0.03	0.05	0.06
	(0.06)	(0.06)	(0.07)	(0.06)
30 or more	0.07	0.05	0.06	0.06
	(0.06)	(0.05)	(0.06)	(0.05)
5-years degree or more	-0.09**	-0.08*	-0.09**	-0.10***
	(0.04)	(0.04)	(0.04)	(0.04)
Taught subjects:				
Reading and Maths	-0.03	0.02	-0.04	-0.04
	(0.08)	(0.08)	(0.08)	(0.08)
Reading	-0.07	-0.01	-0.07	-0.05
	(0.09)	(0.08)	(0.08)	(0.08)
Training:				
Attending courses	-0.02	-0.01	-0.02	-0.01

Dependent variable: Deviation to school-average test scores. Standard errors clustered at the school level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

(*Continued on next page*)

Table 17: (continued)

	Tutor	Students	Tutor	Students
	(1)	(2)	(3)	(4)
	(0.03)	(0.04)	(0.03)	(0.04)
Working teams at school	-0.04	-0.03	-0.04	-0.03
	(0.04)	(0.04)	(0.04)	(0.04)
Congresses and teaching projects	0.02	0.03	0.03	0.04
	(0.03)	(0.03)	(0.03)	(0.03)
Type of warnings to students:				
Letter to the family	0.00	0.00	0.00	0.00
	(0.03)	(0.04)	(0.03)	(0.03)
Temporary class suspension	-0.02	-0.04	-0.02	-0.03
	(0.05)	(0.05)	(0.05)	(0.05)
Person asking for a meeting:				
Parents	-0.02	-0.02	-0.02	-0.02
	(0.04)	(0.05)	(0.05)	(0.04)
Teacher	0.03	0.03	0.04	0.04
	(0.05)	(0.05)	(0.05)	(0.05)
Number of meetings with students' parents	-0.03	-0.03	-0.04	-0.02
	(0.03)	(0.03)	(0.03)	(0.03)
Teacher at 3^{rd} and 4^{th} grades	0.03	0.04	0.04	0.04
	(0.04)	(0.04)	(0.04)	(0.04)
Contextual characteristics	Yes	Yes	Yes	Yes
Observations	12547	11417	12364	11668
R^2	0.08	0.09	0.08	0.10

Dependent variable: Deviation to school-average test scores. Standard errors clustered at the school level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendices

A Sample selection

TO BE COMPLETED.