Teaching Styles and Achievement: Student and Teacher Perspectives*

Ana Hidalgo-Cabrillana
Universidad Autonoma de Madrid

Cristina Lopez-Mayan†
Universitat Autonoma de Barcelona

October 31, 2016

Abstract

We analyze to what extent using different teaching styles is related to achievement. As a novelty, we measure in-class work using information from several sources: teachers and students. We find that teamwork and student discussions (modern practices) are strongly related to better achievement, and individual work and rote learning (traditional practices) are detrimental. The effect of the teaching style is even larger using students' reports. Only with that perspective traditional and modern styles have heterogeneous effects across subgroups, suggesting differences in the perception of teaching styles. Therefore, who reports the class work -teachers or students- matters for the results.

JEL classification: I20; I21; J24

Keywords: Students and teacher reports; Test scores; Teacher quality; Modern and traditional teaching

*We are very grateful to valuable comments by Manuel Arellano, Jordi Caballé, Raquel Carrasco, Jesús Fernández-Huertas Moraga and seminar attendants at Universidad Carlos III de Madrid, Bank of Spain, Universidad Autónoma de Madrid, Centre for Economic Performance (LSE), Universitat de Barcelona, SAEe in Palma de Mallorca, EEA-ESEM Congress and WinE Mentoring Retreat in Toulouse, ESPE Conference in Braga, IAAE Conference in London, Encuentro de Economía Aplicada in Gran Canaria, Jornadas de Economía de la Educación in Valencia, IWAEE in Catanzaro, RWI Research Network Conference on the Economics of Education in Berlin, Annual Meeting of the Portuguese Economic Journal in Azores and EALE Conference in Ghent. We are very grateful to the Instituto Nacional de Evaluación Educativa for helping us with the database. We also acknowledge financial support from the Fundación Ramón Areces grant (*Ayudas a la Investigación en Ciencias Sociales 2012*). All errors are our own responsibility.

†Corresponding author: clopezmayan@gmail.com, Department of Economics, Edificio B, Campus de Bellaterra, 08193 Cerdanyola del Valles, Barcelona (Spain).

1 Introduction

Teacher effectiveness, measured through teacher fixed effects, has a significant impact on student's cognitive achievement (Rockoff (2004), Rivkin et al. (2005), Hanushek and Rivkin (2006), Hanushek (2006, 2011)) and on student's development of non-cognitive skills (Jackson, 2013)[1]. However that finding contrasts with the lack of consistent evidence on the relationship between observed teacher characteristics and student achievement (Hanushek and Rivkin, 2006). Among the exceptions, Rockoff (2004) and Rivkin et al. (2005), which find small effects of the first years of experience, and Dee (2007), which obtains that opposite-gender teachers reduce achievement. A related line of research has found that how teachers actually work in class does a better job than teacher's characteristics to explain student achievement (Van Klaveren (2011), Schwerdt and Wuppermann (2011), Lavy (2016), Bietenbeck (2014))[2]. A potential concern however is that those papers measure teaching practices using the students or the teachers as the unique source of information. Since teaching is a complex process, using only the perspective of one of the agents involved may be problematic if the individuals perceive in-class work differently. In that case, using one or another perspective will not be neutral for the results.[3]

In this paper we use the teacher's and the student's perception of teaching practices in order to analyze to what extent different teaching styles affect student achievement. We distinguish two types of teaching styles: traditional and modern. Broadly speaking, the traditional style is characterized by the use of rote learning and individual work and the modern style by teamwork and involvement of students in discussions and presentations.

We use data from a national assessment program conducted in 2009 in Spain, *"La Evaluación General de Diagnóstico"* (EGD). This program evaluates fourth grade students (nine years old) in mathematics and reading. All the students from each class perform the test and in most schools the EGD evaluates two classes. Classes in fourth grade are organized around the main teacher, the tutor, who teaches most of the subjects, including usually math and reading.[4] Students have the same classmates for the entire school day. The EGD allows linking each student with her tutor and collects rich information about them, including how classes work. The teacher and the students answer the same set of questions about teaching practices. We use those answers to measure the use of traditional and modern teaching styles in class according to the perspective of the teacher and

[1]Hanushek (2011) quantifies that an effective teacher is equivalent to advancing knowledge in one academic year.

[2]Teaching practices affect other important outcomes. Recently Algan et al. (2013) find that the use of a modern teaching style promotes the formation of social capital.

[3]In health economics for instance different observers perceive the same child health differently, leading to different estimates of the income gradient in child health (Johnston et al., 2010).

[4]Throughout the paper, we use the terms "teacher" and "tutor" interchangeably.

the students. We follow the Zemelman et al. (2005)'s taxonomy to classify the teaching practices as traditional or modern. As we discuss later, traditional and modern practices can be complements rather than substitutes and so the aggregate measures do not imply a trade-off between using one or another style.

Our empirical strategy exploits the between-class within-school variation in teaching practices and test scores to identify the effect of different teaching styles on student achievement. This type of analysis is challenging because the non-random allocation of students to schools and to classes within school introduces bias in the estimate of teaching practices. By exploiting within-school variation, we deal with the bias from the endogenous selection between schools. The within-school sorting should not be a major concern since the Spanish schooling system is neither track-based in primary education, nor characterized by the practice of "teacher shopping" by parents. We conduct an exhaustive analysis to show the lack of systematic assignment of teachers and students with certain attributes to the same class. Nevertheless, we also control for a rich set of teacher variables and student characteristics in order to minimize the potential bias due to unobserved traits.

This paper is closely related to the literature identifying the best teaching style.[5] Schwerdt and Wuppermann (2011) and Van Klaveren (2011) study the effect of the percentage of time spent in lecture-style teaching using the *Trends in Mathematics and Science Study* (TIMSS) wave of 2003 for US and Netherlands, respectively. Both papers use a between-subject within-school strategy to control for unobserved student traits. Schwerdt and Wuppermann (2011) find that shifting time from problem solving to lecturing results in an increase in student achievement. This result is in line with Brewer and Goldhaber (1997). However, Van Klaveren (2011) find no relationship between time lecturing and student performance. Lavy (2016) analyzes the effect of traditional and modern teaching on student achievement in Israel using panel data of pupils in fifth and eighth grade. His identification strategy is based on the within-school change in exposure to teaching practices among students attending both grades. Lavy (2016) concludes that traditional and modern practices have a positive effect on test scores and do not necessarily crowd out each other. Bietenbeck (2014) analyzes the effect of traditional and modern teaching practices on math and science test scores using the TIMSS wave of 2007. He estimates a student fixed-effect model, where identification relies on the different student exposure to teaching practices between math and science. Traditional teaching has a positive effect on overall test scores while modern teaching is not significant. After splitting overall scores by cognitive skills, modern practices have a positive effect on reasoning, while traditional

[5]Using an experimental approach, Dobbie and Fryer (2013) and Fryer (2014) find that other school practices -teacher feedback, data-driven instruction, increasing instructional time, high-dosage tutoring and culture of high expectations- increase achievement of students from charter and public schools.

teaching increases knowing and applying skills.

In sum, these studies show that teaching practices matter. In the last years, the proposals to reform education in different countries advocate a greater use of modern teaching practices in detriment of a traditional learning style.[6] But this recommendation contrasts with the still scarce and not conclusive results to identify the best teaching style. This paper provides more evidence on this issue and extends beyond those previous papers in the following.

First, in contrast to previous literature, we estimate the effect of teaching practices using both the perspective of the teacher and her students. Previous works use only one of these perspectives, usually the students, to measure in-class work. Information reported by students and teachers have different advantages and disadvantages (Goe et al., 2008). Students' reports about teaching are useful because they provide the perspective of students, the recipients of the teaching practices. However, student responses are subject to bias. Students do not know all the aspects of teaching. Pupils may also answer about in-class work influenced by personality characteristics of the teacher or by their grades. In contrast, teacher self-reports have the advantage that teachers know their own abilities, the class context, and how they work in class. However, teacher responses are also subject to potential biases. Teachers may misreport their practices to adjust them to the "social desired" practices or because they believe that they are applying a certain practice when actually they are not. Therefore, since both student's and teacher's responses on teaching practices are self-reported measures with different potential reporting bias, using both sources of information will improve our understanding of the role of teaching practices on student achievement. Goe et al. (2008) recommend assessing teacher effectiveness gathering data from more than one source, especially if one of these sources are students' reports.

Second, we analyze the effect of teaching practices on performance of younger students. Previous papers analyze that effect for students at eight grade. Research on early development outcomes -for instance Heckman (2008)- highlights the importance of understanding at the earliest stages how the education process successfully improves student achievement in order to prevent future dropouts and improve outcomes later in life (Chetty et al., 2011).

Finally, none of the previous studies has analyzed the impact of the teacher characteristics and the teaching practices on the achievement of Spanish students. However, it is essential to provide evidence on the role of the teacher because the Spanish educational system faces serious problems, such as the high dropout rate (23.5% in 2013 according to Eurostat, far away from the 10% target of Europe 2020 strategy) and the lack of excellence

[6]For example, the recent reforms in Finland (http://www.minedu.fi/OPM/Verkkouutiset/2015/03/curricula.html?lang=en) and UK (https://www.gov.uk/government/news/new-curriculum-will-make-education-system-envy-of-the-world)

(as shown by the low performance of Spanish students in PISA). Although these problems are measured at the end of secondary schooling, they may come from earlier educational stages. Education is a cumulative process and thus it is important to analyze how schooling inputs affect learning outcomes in primary education.

We find that the use of modern practices is related to better student achievement, while the traditional teaching if any is detrimental. The magnitude of the effect is larger when the teaching style is measured using the practices reported by the students. The effect of the practices is heterogeneous across subjects. Traditional teaching is detrimental for reading scores but only when the students report the practices. Modern teaching is positively related to scores in reading and math, although the latter effect only arises if we use the students' answers. When the practices are reported by the teacher, the estimates hardly differ for boys and girls, students from public and private schools, and students with the same or different tutor in the previous grade. When the practices are reported by the students, the effects are different. Boys do no benefit from using any particular teaching style, while girls gain from modern practices and lose from traditional ones. Modern and traditional practices are related respectively to higher and lower scores in public schools but not in private ones. Students with the same tutor in the previous grade seem to benefit less from the modern style than students with a different tutor. The fact that the heterogeneous effects arise when the practices are reported by the students but not by the tutor suggests that the differences in the estimates may result from a different perception of the teaching styles.

In line with previous literature, pupils' achievement is not correlated to teacher's gender or experience. However, having a teacher with more than three years of college is negatively correlated with achievement, suggesting a pattern of negative selection of those teachers into primary education in Spain.

The rest of the paper is organized as follows. Section 2 describes the database and explains the construction of the teaching measures. Section 3 explains the empirical strategy. Section 4 presents the results and the sensitivity analysis. Section 5 concludes.

2 Data

We use data from *"La Evaluación General de Diagnóstico"*, the national assessment program conducted in 2009 by the *Instituto Nacional de Evaluación Educativa* (INEE), the public institution for the evaluation of the Spanish education system. The EGD evaluates the competencies of fourth-grade students in mathematics and reading using a standardized test following the PISA methodology. In reading, the EGD evaluates the competencies in understanding texts and the literacy to write the own ideas and in mathematics, it assesses

the ability to do basic math and to apply the mathematical reasoning in solving problems.

The EGD evaluates 28,708 pupils from 900 schools selected with a two-stage stratified sampling method to ensure that the results are representative both at the national and regional level, and for public and private schools. In the first stage, for each stratum, schools are selected with probabilities proportional to their fourth grade enrollment. In the second stage, one or two fourth grade classes are randomly sampled. All the students from the selected classes are evaluated but the tests of pupils in special situations are not included in the final EGD database (students with serious special needs or immigrant students who recently entered the Spanish schooling system). The EGD scales test scores so that the average in each domain (mathematics and reading) is 500 and the standard deviation is 100. In order to interpret coefficients as fractions of a standard deviation we standardize scores to have a mean of zero and standard deviation of one.

The EGD does not provide actual test scores. The student's overall achievement is available through five *plausible values*, or imputed values. Providing plausible values instead of actual test scores is a standard practice in international assessment programs (*Programme for International Student Assessment*, TIMSS). In this type of evaluations, each student answers a limited number of test questions. Then, those answers and the student's family background are used to estimate the proficiency distribution of the student by applying the Item Response Theory. Plausible values then are random draws from that distribution.

The EGD includes rich information about students and about the tutor of each fourth-grade class. The tutor teaches most subjects, including usually the core ones -math and reading. Students have the same classmates for the entire school day. Indeed, students are assigned to a class in first grade and they usually continue with the same classmates until sixth grade, the last grade of primary education. The whole fourth-grade class spends a large fraction of the school day with the tutor.

The tutor is also the person who follows the performance of the students, monitors the class climate and meets with parents. Aside from the relatively standard set of tutor characteristics, the EGD provides information about (i) the teaching practices used in class; (ii) whether the tutor teaches both math and reading, only math, only reading, or none; and (iii) the tutorial work, such as the number of meetings with parents and whether the tutor was the tutor of the class in third grade as well.

The original sample contains 28,708 pupils, 1,358 classes and tutors -since there is one different tutor per class- and 900 schools. From this sample, we drop (i) students with missing math or reading scores; (ii) classes with an extremely small number of students (less than five); (iii) students who take the test but did not fill in the questionnaire or whose tutor did not answer her questionnaire; (iv) classes whose tutor does not teach math or reading to ensure that teachers in the final sample are the instructors of at least one of

the analyzed subjects; and (v) students and teachers with missing information in basic variables such as gender, parents' education, years of experience, and teaching practices. As we discuss later, our identification strategy relies on within-school variation, so we also rule out the schools with only one sampled class. The final sample contains 12,113 students, 736 classes/tutors and 368 schools. Of the 368 schools, 69% are public and the remaining 31% are private and semi-private (private schools publicly funded). Despite the reduction in size, the final sample is not significantly biased with respect to the initial one and is still representative of the target population of fourth-grade students in Spain as Tables A.1 and A.2 in the Appendix show.

Table 1 presents the characteristics of the fourth grade tutors in primary school in Spain. They are mostly women, have more than twenty years of experience, teach both mathematics and reading, and were tutor of the same class in third grade. Regarding educational qualification, 17% of the tutors hold more than a three-years university degree (five-years degree, master or PhD), which is the minimum educational level required by Spanish authorities to teach in primary education. Tutors meet with parents an average of three times per school year -presumably, once per quarter- and it is more usual that the tutor asks for the meeting.

The average number of pupils per class is 21.14 with a standard deviation of 4.85. We compute the class size as the total number of surveyed students in the class in the initial sample.

Table 2 reports students' characteristics. Half of fourth-grade pupils are girls and 6% has repeated at least once. The proportion of non-Spanish pupils is 7%, coming mainly from Latin America, Non-Western Europe and Morocco. Most students started school at three years old or less, which is the usual age to begin school in Spain. Regarding household composition, 7% of the students live with a single parent and 85% live with at least one sibling. Although mothers are slightly more educated than fathers they are unemployed or inactive in a higher proportion than fathers[7].

Table 3 shows that on average girls perform better than boys in reading, while boys perform better in math. Students in private schools perform better on average in both subjects than students in public schools.

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

2.1 Teaching practices

The information about teaching practices is derived from the question, "How often do you use the following teaching practices in your lessons this school year?". On a point-four scale, possible answers are "Never or almost never", "Sometimes", "Almost always", and "Always". Teachers respond about each of the following practices: (a) "Most of the time I teach by telling", (b) "Students present works or topics to classmates", (c) "While I teach, I ask students questions about the lesson", (d) "While I teach, students ask me doubts", (e) "I promote discussions", (f) "Students work on exercises and activities proposed by me", (g) "Students work individually", (h) "Students work in small groups", (i) "I give different exercises or activities to best/worst students". We do not consider this last item because it reflects the level of students in class and it would lead to a problem of reverse causality in the estimation. According to the taxonomy by Zemelman et al. (2005), practices (b), (e), and (h) are classified as modern, and practices (a), (f), and (g) as traditional (Table 4). It is not possible to unambiguously match items (c) and (d) as traditional or modern. In principle, item (c) may be thought as traditional and item (d) as modern, but classifying (d) as traditional and (c) as modern is reasonable as well.

The EGD data supports the theoretical classification derived from the Zemelman et al. (2005)'s taxonomy. Table 5 shows the correlation among the tutor's answers to all items. Modern items (b), (e), and (h) are positively correlated with coefficients around 0.26. The same pattern appears for traditional items (a), (f), and (g), with coefficients ranging from 0.13 to 0.30. Items (d) and (c), classified as modern and traditional in this Table, are positively correlated with the respective modern and traditional items. At the same time, (d) is positively correlated with traditional items, and (c) with modern ones (see the bottom left of Table 5). This clear pattern of positive cross-correlations does not appear for the rest of items. Moreover, the correlation between (c) and (d) is quite high (0.47). Therefore, we exclude items (c) and (d) from the baseline measure of teaching practices. In Section 4.3 we check the robustness of the results to include those items.

Following Lavy (2016) and Bietenbeck (2014), for the ease of interpretation, we rescale the answers to each item by assigning a numerical value as follows: 0 to "Never or almost never", 0.33 to "Sometimes", 0.67 to "Almost always", and 1 to "Always". Thus, like in previous works, responses are interpreted as the proportion of the time used in that practice. The aggregate measure of traditional teaching practices is the mean of the teacher's answers to items (a), (f) and (g); and the aggregate measure of modern teaching practices is the mean of the teacher's answers to items (b), (e) and (h). Table 6 shows that teachers report that they use traditional and modern practices, respectively, 66% and 43% of the class time on average.

We should note that both indexes do not imply a trade-off between using one or another style in class. Teaching practices can be complements rather than substitutes (De Witte and Van Klaveren, 2014). For instance, one possible activity proposed by the teacher (item (f), traditional) may be to promote discussions in class (item (e), modern). Indeed, Table 5 shows a positive correlation between these two items. Moreover, the question about in-class work does not impose any restriction on the complementarity or substitutability among the practices because the answer on the frequency of use for one practice does not restrict the answers for the rest of practices. We thus do not restrict the aggregate measures either. This explains that the means of traditional and modern indexes sum above one for each respondent (Table 6). In the estimation we include jointly the traditional and modern indexes and the estimated coefficient of one of the indexes should be interpreted as the effect on test scores holding constant the other one.[8] We then assess the sensitivity of the results to restrict that the total proportion of time using all practices must be equal to one or, in other words, that all teaching practices are substitutes.

Table 7 shows that the correlation between the traditional and modern indexes is zero. This may be explained by the opposite cross-correlations among individual items (see the bottom left of Table 5).

The EGD survey also asks students about teaching practices. The question is "In general, how is in-class work?". The list of practices provided to the students correspond exactly with items (a) to (h) from the teacher questionnaire. Students answer using the same four-point scale explained above. We also rescale the students' answers to each item by assigning a proportional value from zero to one. We construct the modern and traditional indexes by averaging the responses of the students in the class -excluding the student's own response- to the modern and traditional items.

Using the students' or the teacher's reports about teaching practices to estimate their effect on test scores may be subject to different biases or measurement errors. In principle, teachers' reports may seem more accurate because teachers know their own abilities, the class context, and how they work in class, while the students do not know all the aspects of teaching. In addition, the question to the students is about class work in general and, although the tutor teaches most subjects, we cannot disregard that some students might answer thinking on another teacher. Pupils may also answer influenced by personality characteristics of the teacher or by their grades. However, teachers may misreport their practices either intentionally (to adjust them to the "social desired" teaching) or unintentionally (believing that they are using a certain practice when they are actually not). Since it is not clear whether the students' or the teachers' answers are less subject to bi-

[8] Lavy (2016) and Bietenbeck (2014) also use this approach.

ases, unlike previous works, we use both sources of information to estimate the relationship between the teaching style and achievement.

Table 6 shows that the pupils' answers about modern teaching are on average close to the tutors' responses while students report a traditional index higher than teachers. To gain further insight about to what extent student and tutor perspectives differ, Figure 1 shows the distributions of the differences in the indexes between tutor and students and Table 8 the descriptive statistics. The indexes constructed using the student perspective are the average of the answers of all students in the class. Figure 1 demonstrates clearly the differences between tutor and students. Students tend to report a higher use of traditional practices than tutors (distribution shifted to the left). In the modern index, the differences exist but overall they almost compensate each other. The dissimilarity of tutor and students reports is also evident in the correlation between respondent's indexes, shown in Table 7, where cross-respondents correlations are just 0.20 and 0.25 for the traditional and modern index respectively.

3 Empirical Strategy

The effect of the use of traditional and modern teaching styles on student achievement can be estimated using the following education production function:

$$y_{ics} = \alpha + \gamma' TP_{cs} + \lambda' T_{cs} + \beta' X_{ics} + \phi_s + \varepsilon_{ics} \tag{1}$$

where y_{ics} is the standardized test score of student i in class c at school s, TP_{cs} is the vector of traditional and modern teaching indexes, T_{cs} is the vector of tutor variables and class size, X_{ics} is the vector of student characteristics, ϕ_s is the school fixed effect and ε_{ics} is the error term. We run separate regressions for the indexes constructed using the tutor's and the students' answers. When we use the latter, the indexes are the average of the responses of students in class c excluding the student's own response.

The identifying assumption of the effect of traditional and modern teaching on student achievement (γ) is that teaching practices are uncorrelated with the error term conditional on the other regressors. One of the potential confounding factors is the between-school sorting, that is, the endogenous selection of students across schools. This sorting arises because parents do not choose the school randomly. They may prefer a school where students share a similar socioeconomic background, that hires teachers with some specific characteristics or that has certain teaching philosophy. To account for between-school sorting, we focus on the schools with two sampled classes and include school fixed effects in (1). In this approach the identification of γ relies on the variation in teaching styles and test scores across the two classes of the school. As shown in Table 9 most of the variation

arises between schools but still a significant amount of the variation in the teaching indexes -one third- happens within a school.

Even after controlling for between-school sorting, the error term may still include unobserved traits of the student (μ_{ics}) and the teacher (η_{cs}) that may bias the estimate of γ. In particular, γ would be biased:

(i) If there are student unobserved traits that have a direct effect on y_{ics} while they are correlated with the teaching style, i.e., $corr(\mu_{ics}, TP_{cs}) \neq 0$. This would happen if students are not randomly allocated 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 (reverse causality). For example, if high-ability students are assigned to the same class and the teacher decides to use a certain teaching style with that class, the estimate of γ will be biased. It is important to note that although μ_{ics} affects scores, if students are randomly assigned to classes and teachers do not adapt their teaching style to the ability of the class, γ will not be biased.

(ii) If there are unobserved teacher traits that have a direct impact on y_{ics} while they are correlated with the teaching style, i.e., $corr(\eta_{cs}, TP_{cs}) \neq 0$. This would happen if the unobserved ability or motivation of the teacher affects the choice of the teaching style, and it has also a direct effect on student test scores, aside from the effect through the teaching practices.

3.1 Within-school selection of students, teachers and teaching style

As we discuss above the non-random assignment of students or teachers to classes may lead to a biased estimate of the effect of the teaching style. Possible sources of this within-school selection are track-based education and "teacher shopping", i.e., parents "buy" the teacher assigned to their children. They are not a concern in our analysis because in Spain "teacher shopping" is absent or very rare and primary education is neither track based nor has an explicit rule to assign students to classes. However, we cannot fully disregard the lack of within-school sorting since the school principal still may decide to assign students and teachers to classes following a non-random rule. To assess whether students are not randomly allocated to teachers and whether teachers adapt the teaching style to their students or to their own characteristics, we conduct the following analysis.

First, we investigate to what extent students with certain family characteristics are more likely to be in classes with certain type of teacher. To this end, we regress a set of teacher variables on the socioeconomic characteristics of the class:

$$t_{cs} = \alpha_0 + \alpha_1' X_{cs} + \phi_s + v_{ics} \tag{2}$$

where t_{cs} is a characteristic of the tutor of class c in school s; ϕ_s is the school fixed effect; and X_{cs} is the vector of socioeconomic characteristics of class c at school s, which are the average of these student characteristics: female, living with siblings, living in single-parent household, non-Spanish origin, repeater, parents' education and parents' labor status. Table 10 reports the results. Each column represents a separate regression. We do not find a systematic within-school relationship between the teacher characteristics and the socioeconomic background of the class. Note especially the lack of relationship with the average parental human capital. The bottom of the Table shows the F-test of the joint significance of the class characteristics. F-statistics do not reject the null hypothesis that the joint effect is zero. Only for the tutors who hold a five years degree or more is possible to reject that hypothesis but at ten percent level. However, no clear pattern emerges regarding family background. Those tutors tend to be in classes with a higher proportion of non-Spanish students but also with a lower proportion of students with unemployed fathers.

Second, we analyze whether teachers use different teaching styles when face classes with different socio-economic background. Table 11 shows the results from the estimation of regression (2) where the dependent variable is one of the teaching indexes. We do not find that the socioeconomic characteristics of the class are related to the use of certain teaching styles. F-statistics do not allow rejecting the null hypothesis that the joint effect of the class characteristics is zero. It is especially useful to demonstrate the lack of pattern of selection based on parental human capital variables. Some class characteristics are individually correlated with the teaching style but without showing a clear relationship between the use of traditional or modern practices and the socioeconomic background of the students. Moreover, the positive relationship between a higher use of the modern style and classes with a larger proportion of students repeating, with siblings or born outside Spain may reflect student reporting biases instead of an endogenous selection of teaching practices because those correlations do not appear when the tutor reports the practices.

Third, even though classes are formed more or less randomly, they may receive other school resources differently. For instance, a teacher with a specific teaching style may be assigned to classes of certain size. To check this, we run

$$tp_{cs} = \lambda_0 + \lambda_1 size + \lambda_2 size^2 + \phi_s + \varsigma_{ics} \tag{3}$$

where tp_{cs} denotes teaching practices (traditional or modern); $size$ is class size and ϕ_s is the school fixed effect. Results are shown in Table 12, where each column represents a separate regression. Columns one and three do not include school fixed effects. Using the tutor's answers, class size and class size squared are not significantly correlated with the teaching style, nor individually neither jointly, once we account for the school fixed effect[9].

[9]We repeat the analysis using the students' responses and the results are the same. The Table is

Finally, we regress the traditional and modern indexes on tutor variables, class size and school fixed effects:

$$tp_{cs} = \theta_0 + \theta_1' T_{cs} + \theta_2 size + \phi_s + \psi_{ics} \tag{4}$$

The purpose is to check whether using a teaching style is correlated with certain teacher characteristics after controlling for school and class size. We estimate regression (4) for the teaching indexes derived from the tutor's and students' answers. Table 13 shows the results. F-statistics do not allow rejecting the null hypothesis that the joint effect of the teacher characteristics and class size on the use of traditional and modern teaching is zero. When the tutor reports the practices, only holding a five-years college degree and having thirty or more years of experience are significantly correlated at ten percent level with the use of a traditional style. Modern practices are not significantly associated with any tutor characteristics. It is noteworthy the lack of pattern of correlation with the years of experience, although the relatively homogeneous teacher profile shown in Table 1 (mostly women with more than twenty years of experience) could explain it. When the students report the practices, most characteristics are not significantly related to the teaching style either. Being a tutor female is associated with a lower use of the traditional style and having more than twenty-five years of experience is negatively correlated with the use of modern teaching. Being the tutor in third grade is related to an increase of the same magnitude in the use of both styles. Since this relationship only arises using the students' answers, we cannot disregard that it reflects some type of reporting bias -maybe students with the same tutor in third and fourth grades report better the teaching style than those with a different tutor- instead of reflecting sorting of teaching styles.

Previous findings do not support that students and teachers are systematically assigned to classes or that teachers with certain characteristics self-select into a specific teaching style. Nevertheless, we cannot fully disregard the presence of unobserved traits that may bias the estimates. In order to minimize that potential bias, we consider a broad set of student and teacher variables for the estimation of regression (1). The vector of student characteristics (X_{ics}) includes gender, country of origin, repeater, mother and father's education, mother and father's labor status, living in single-parent household, living with siblings, born in the fourth quarter, age at starting school, and whether a private tutor or someone in the family helps the student with homework. Note that this set of controls includes variables as proxy for unobserved student ability and previous performance (for instance, repeater and parents' education). The vector of tutor characteristics (T_{cs}) includes the standard controls used in the literature -gender, experience, type of degree, class size-, and additional variables measuring the tutorial work (number of meetings with parents,

available upon request.

whether the tutor or the parents ask for a meeting, being tutor of the class in third and fourth grades) and whether the tutor is the instructor of math and reading or only one of them. All those variables also capture partially the teacher unobserved effort and ability.

In sum, although we cannot rule-out completely the presence of unobserved teacher or student traits, and consequently we have to be cautious in interpreting our estimates as causal, we should note that (i) we conduct an exhaustive analysis showing no evidence of within-school sorting; (ii) we do not find evidence of the correlation of teaching practices with observed teacher and class characteristics, so it is plausible to assume that selection on unobservables is neither a big concern; (iii) we include a broad set of regressors to control for potential within-school differences in the student background and the tutor across classes. Finally, note that the potential problem of endogeneity of the teaching practices with the test scores in a particular subject is ameliorated because the tutor and students answer about the practices generally used in class, instead of about the particular teaching style used in math and reading.

4 Results

Regression (1) is run five times -one for each plausible value- using the student weights provided by the EGD. All the tables report the average estimates and the mean R-squared. Standard errors are adjusted upward for the imputation variance -i. e., the dispersion introduced because of using imputed values- following the formula provided on page 118 in OECD (2009). Standard errors are additionally clustered at the class level[10].

Table 14 presents the results from the estimation of specification (1) jointly for test scores in math and reading adding a dummy variable for math. Columns (1) to (3) show the results corresponding to the estimation using the teaching practices reported by the tutor, and columns (4) to (6) the results using the practices reported by the students. In columns (1) and (4), we estimate including only the modern and traditional indexes and the dummy for math. In columns (2) and (5), we add class size and teacher characteristics. In columns (3) and (6), we include student characteristics[11]. The coefficients of interest do not change much when introducing teacher and student characteristics. This shows that within-school sorting is not driving the results, as supported by the evidence obtained in previous section. When the teaching practices are reported by the tutor, student achievement is not correlated with the traditional style and is positively correlated with the modern teaching. However, the coefficients are measured imprecisely and the estimates are not significant. When the teaching practices are reported by the students, the estimates are larger. In the

[10]For the estimation we use the Stata command *pvcluster* developed by Bietenbeck (2014).

[11]Table A.3 in Appendix shows the estimates of the student characteristics.

specification involving all the controls, a 10% increase in the use of modern practices is associated with a 3.4% of a standard deviation increase in test scores and a 10% increase in the use of traditional practices is related to a 2.6% of a standard deviation decrease in test scores. Note the magnitude of the estimates. They are larger than the estimates of any other characteristic of the tutor. Compared to students characteristics in Table A.3, the magnitude of the modern teaching is similar to having parents with university education. The traditional style is among the largest negative estimates, after being repeater and starting school older than four years old.

In line with previous literature, we do not find strong evidence that the student achievement is correlated to observable teacher characteristics. The effect of being female is negligible and not significant. The years of experience are associated with higher test scores, although the estimates are small and barely significant. Student achievement is lower if the tutor teaches reading but not math and is higher if the students had the same tutor in third grade. Surprisingly those teachers who hold a college degree of five years or more are associated to a lower student achievement of 0.07 standard deviations compared to the teachers with a three-years college degree. The estimate is significant at one percent level and is robust to adding controls and using the tutor's or students' answers for the teaching practices. Since holding the three-years college degree is the minimum requirement to teach in primary education in Spain, the negative estimate may suggest that teachers with more years of college are negatively self-selected. That is, maybe those teachers became primary education teachers once they did not find a job in the private sector or in secondary education (where the requirement is to hold at least a five-years college degree). Consequently, those teachers may lack motivation or adequate teaching skills, and this would explain the negative estimate that we find. Unlike us, previous works obtain that teachers with more years of education are related to better student performance.

Our results provide new insights on the effect of the teaching style on achievement. Lavy (2016) finds that both traditional and modern teaching practices have positive effects on test scores, larger for traditional teaching. Bietenbeck (2014) concludes that only traditional practices have a statistically significant and positive effect on overall test scores. Schwerdt and Wuppermann (2011) find that spending more time lecturing is associated with higher test scores. In contrast, we obtain that the use of modern practices is related to better student achievement, while traditional teaching is detrimental. Several reasons may explain these differences. We analyze younger students from a different country, and the particular teaching practices collected by the EGD are slightly different than those used by other authors. Moreover, previous results are found using one source of information (teachers in Schwerdt and Wuppermann (2011), and students in Lavy (2016) and Bietenbeck (2014)). Our findings show that using the tutor or the students as source of

information about in-class work leads to different estimates.

4.1 Heterogeneous effects

Previous results are obtained assuming that the effect of modern and traditional practices is the same in math and reading. Now, we explore whether the effects are different estimating specification (1) separately for each subject. Alternatively we could estimate a specification including the interaction terms of math and reading with the teaching indexes. However, in that approach the estimates of the interaction terms would capture differences in the effect of practices by subject as well as differences in the effect of other variables. Although the separate estimation implies smaller sample sizes, we prefer this approach to allow the effect of the rest of variables to differ by subject. Table 15 presents the results for the specification including all control variables. When the tutor reports the teaching practices, the traditional and modern styles have a different impact on the performance in math and reading, although only the positive relationship of the modern teaching with reading scores is large and measured precisely. A 10% increase in the use of modern practices is associated with 3% of a standard deviation increase in reading test scores. When the students report the practices, the modern style is related to higher scores in both subjects, although the effect is larger for reading (4.1%) than for math (2.7%). A 10% increase in the use of traditional practices reduces scores by 2.6% of a standard deviation but the estimate is significant only for reading.

Compared with the results in Table 14, math scores benefit from having a tutor with an intermediate level of experience, while reading scores benefit from having either the least or the most experienced tutors. The negative relationship between achievement and having a tutor with more years of education appears in both subjects with a similar magnitude. Compared with a tutor who only teaches math, a tutor who teaches only reading is detrimental for math scores while she does not affect performance in reading.

Table A.3 in Appendix shows that the estimates of the student characteristics present some differences across subjects. It arises the usual gender gap: girls obtain higher scores in reading and lower in math than boys. The older the student started at school, the lower the scores, especially in reading. Latin American students obtain the lowest scores in math and reading, even though Spanish is their mother tongue. Parents' education is positively correlated with scores and the association is larger for reading. It is interesting to mention that the parents' labor status is not correlated with scores except if the mother is employee. The positive association may suggest that employee mothers spend less time with children than inactive and unemployed mothers but the time spent is of higher quality.

In Table 16 we explore whether the effect of the teaching style is different for top and

bottom performers. Separately for math and reading, we define the top performers as the students who score above the 75 percentile of the score distribution and the bottom performers as the students who score below the 25 percentile.[12] Using the tutor's answers, students from the top part of the distribution in general benefit from both teaching styles in math and reading. In contrast, if the students report the practices, top performers lose both from traditional and modern styles in math. Students from the bottom part of the distribution lose especially from the use of modern practices in math. In reading they lose from the traditional style and gain from the modern one, particularly if we use the students' answers. However, standard errors are quite large and the estimates are in general not significantly different from zero.

To gain deeper knowledge about the effect of the teaching practices on achievement, we explore whether there are differences by the gender of students, the type of school and the type of tutor. By type of tutor we refer to whether the students had or not the same tutor in third and fourth grades. Each panel in Table 17 shows the results of estimating for math and reading jointly and separately. In Panel A the estimated coefficients of modern and traditional teaching do not reveal differences across male and female students if the tutor reports the practices: the modern teaching has the same positive and significant effect for reading scores of boys and girls. However, if we use the students' reports, striking differences appear. For girls, the use of traditional practices is strongly associated to lower scores in math and reading, while the effect for boys is negligible and not significant. In contrast, girls' reading scores benefit from using the modern style while boys' scores do not. Nor boys' neither girls' math scores are significantly correlated with using modern teaching practices.

In Panel B we stratify the sample by public and private schools. Using the tutor's answers, no significant differences appear except the positive relationship between reading scores and modern teaching for students in public schools. With the students' answers, the use of modern practices is significantly associated to higher math and reading scores in public schools but not in private ones. Traditional teaching decreases reading scores but only among students from public schools.

In Panel C we split the sample among the students who had the same tutor in third and fourth grades and the students who had a different tutor. Table 1 shows that 74% of the tutors were also the tutors of the class in third grade. According to the practices reported by the teacher, the modern teaching is significantly related to higher scores in reading for the students with the same tutor in third and fourth grades. For students with a different tutor, the estimate is also positive but it is not significantly different from zero. When the

[12] We use the first plausible value to obtain the score distribution.

teaching practices are reported by the students, several differences arise. The modern style is positively related to math and reading scores but with a lower magnitude for students with the same tutor in third and fourth grade. The traditional style is strongly related to lower scores in reading but only for students who did not have the same tutor in the previous grade.

4.2 Differences in students' and teachers' responses

From previous findings we conclude that the effect of the teaching style is heterogeneous across subjects and several subsamples mainly if the teaching practices are reported by the students. This suggests that the heterogeneous effects may not be the consequence of differences in the effect of the teaching style. Instead, they may partly result from a different perception of the traditional and modern styles for instance by boys and girls, or students from public and private schools.

First, to explore whether observed characteristics lead to differences in the teaching practices reported by students and teachers, we estimate the following model:

$$TP_{jcs}^{Tut} - TP_{jcs}^{Stu} = \gamma_0 + \gamma_1' T_{cs} + \gamma_2' X_{cs} + \phi_s + public_s + \varepsilon_{ics} \tag{5}$$

where TP_{jcs}^{Tut} and TP_{jcs}^{Stu} are the teaching practices indexes for the style $j = \{$modern, traditional$\}$ as reported by the tutor and the students, T_{cs} is the vector of tutor characteristics and class size, X_{cs} is the vector of the average characteristics of students from class c at school s, $public_s$ is a dummy variable for public school and ϕ_s is a school fixed effect. X_{cs} includes the average scores in math and reading to explore whether differences in the perception of the practices are related to student ability -for instance, high achievers may perceive class work more accurately and respond more similar to her teacher. Unexpectedly, tutor and average class characteristics are not significantly correlated to the difference in modern and traditional indexes. Only ability explains differences in the indexes: students from classes with higher average scores in reading report a lower use of the modern teaching than the tutor. For the sake of brevity we do not show this table but it is available upon request.

Second, to explore whether differences in students' responses are explained by certain observed characteristics, we estimate the following model:

$$ModP_{ics}^{Stu} - TradP_{ics}^{Stu} = \beta_0 + \beta_1' X_{ics} + \beta_2' T_{cs} + \beta_3' X_{cs-i} + \phi_s + public_s + \varepsilon_{ics} \tag{6}$$

where $ModP_{ics}^{Stu}$ and $TradP_{ics}^{Stu}$ are the modern and traditional indexes reported by student i, X_{ics} is the vector of individual characteristics, including reading and math test scores, and X_{cs-i} is the vector of average class characteristics and test scores excluding the student's

own value. We use the first plausible value for test scores but results do not change using the other ones. Table 18 shows the results. Girls and students with high-performing classmates tend to overreport the use of traditional practices. Repeaters, students with higher scores in reading and from public schools tend to overreport the use of the modern style. This may suggest that the significant coefficient of the effect of modern practices on reading could partly come from the reporting bias of students with higher reading scores. Unfortunately, due to the lack of appropriate data it is difficult to provide a clear interpretation of the mechanisms explaining those differences in the perception of the use of modern and traditional practices.

4.3 Sensitivity analysis

In this Section, we conduct several sensitivity tests in order to address potential reservations about our findings. First, the baseline specification includes the traditional and modern indexes jointly because they do not imply a trade-off between using traditional or modern methods in class. A possible concern is whether including jointly the indexes creates a collinearity problem that may influence the results because some traditional items are correlated -although weakly- with some modern items (see Table 5). Table 19 shows that collinearity is not an issue because the results do not change if the indexes are included separately in the regression.

Second, we analyze whether the results hold after considering alternative ways of measuring teaching practices. Instead of aggregating the items, in Table 20 we estimate a specification that includes the six teaching practices. We do not observe any particular item leading our main findings. The baseline estimates are the result of individual effects that either compensate or reinforce each other.

In Panel B of Table 21 we assess the sensitivity of the results to include items (c) "While I teach, I ask students questions about the lesson", and (d) "While I teach, students ask me doubts". The baseline indexes do not include those items because, as we discuss, its classification as traditional or modern is ambiguous. In Panel B1 we redefine the traditional and modern indexes including (c) as traditional and (d) as modern, while in Panel B2, we consider (c) as modern and (d) as traditional. The new estimates are consistent with the baseline results. The estimated coefficients do not change too much but if any they move slightly towards a higher positive effect of the modern style and a lower negative effect of the traditional practices. Then the baseline estimates, obtained without including items (c) and (d), may be seen as conservative or lower bound estimates of the relationship between teaching styles and student achievement.

In Panel C of Table 21 we construct a new measure of teaching practices that restricts

that the total class time allocated to the six traditional and modern practices listed in Table 4 must sum to one. In the baseline indexes, we rescale the answers to each practice from zero to one in order to interpret the responses as the proportion of the time used in the practice. Without imposing any restriction on the total time engaged in teaching practices, the traditional and modern indexes from a class can sum more than one. Indeed, according to the teacher's reports on practices, this happens in 74% of the classes. As discussed in Section 2.1, in the baseline measures we do not impose that the proportion of the time using modern and traditional practices is equal to one because different practices may be complements rather than substitutes. However, measurement error or careless responses (for instance, answering "always" to all items) may also result in traditional and modern indexes that sum above one. In order to assess whether this concern affects the results, we rescale the answers of each individual such as they sum to one. That is, for each individual we aggregate the numerical values assigned to the answers of the six items in Table 4 and weight each answer by the inverse of that sum. In this way, we keep the relative frequency in the use of practices while imposing that the proportion of the time using traditional and modern practices fulfills the time budget constraint. In other words, we impose that all practices are substitutes. The restricted share of the class time using modern practices is then the sum of the weighted time allocated to the three modern items. The proportion using traditional practices is the remaining weighted time and so the new specification includes only one of the measures. Note that these new measures of the teaching style are more restrictive than the baseline measures. Panel C shows that the estimates of the share of time using modern practices are similar to the baseline estimates. We can conclude that measurement error or careless responses about in-class work are not a concern for our findings.

Third, we analyze whether the results are robust to adding the class average of all student characteristics -excluding the student's own value- to the baseline specification[13]. Panel D of Table 21 shows that controlling for those additional variables hardly changes the effect of the teaching style on student achievement. This supports our previous evidence that within-school sorting is not a big concern. If the main results were driven by this type of selection, controlling for sociodemographic characteristics of the class would lead to different results.

Finally, we may think about using the teaching indexes reported by the students as instrument for the indexes reported by the teacher to correct for potential measurement error bias. Note that this approach would not correct for the possible bias introduced by the endogenous selection of the teaching style. However, the first stage estimates show

[13]Regarding the student's origin we control for the percentage of non-Spanish students in the class.

that the instruments behave weakly (results are available upon request). The correlations among the indexes reported by the students and the teachers is not very large, although they are not so low to flag a weak-instruments problem (see Table 7). However, both the F-statistics for the joint significance of the instruments and the extremely low values of the partial R^2 point to the weakness of the instruments.[14]

5 Conclusions

We analyze to what extent using traditional or modern teaching styles in class is related to the student achievement in math and reading in primary education. As a novelty, we measure in-class work using two different sources of information -students and tutor. To deal with the non-random assignment of teachers and students to schools, our identification strategy relies on between-class within-school variation of teaching styles. We show that the systematic selection of students and teachers within-school is not a concern.

We obtain that the use of modern practices is related to better student achievement, while the use of traditional practices is detrimental. The magnitude of the coefficients is larger when the teaching practices are reported by the students. The modern style leads to higher scores in reading and math, but the latter estimate is significant only when the students report the practices. The effect of the teaching style is different for boys and girls, students from public and private schools, top and bottom performers and by type of tutor but mainly when the practices are reported by the students. Results are robust to considering alternative definitions of the teaching practices and to controlling for the average socioeconomic characteristics of the class.

Pupils' achievement is not correlated to teacher's gender or experience, but it is negatively correlated with the teacher's degree. We discuss that this may reflect a pattern of negative selection into primary education of the teachers holding a five-years degree or more. Spanish educational authorities should take into account this misallocation problem when they establish the degree requirements to teach in primary education.

Our findings are important from a policy perspective. They highlight that using the teacher's or students' reports may lead to different conclusions about the relationship between achievement and teaching style. We advocate that future research should use more than one source of information, not only teachers or students. Ideally research might combine those subjective measures with more objective assessment of class work, such as direct class observation by a third person. Direct observation not only may provide reliable information on the teaching practices used but also an alternative to test-based

[14] If the partial R^2 is much smaller than the R^2, the instruments are weak because they are adding little extra to explaining the endogenous regressors after accounting for the rest of variables.

measures to evaluate teacher effectiveness. Test scores do not capture all aspects of student learning. Test-based measures can identify the practices more effective to increase test skills in detriment of other practices not so effective for the test but that may give more valuable skills to the students. Related to this, math and reading questions may capture different skills and this also may explain why the effect of teaching styles is different across subjects. A more accurate understanding of the relationship between in-class work and student outcomes is necessary to identify the most effective practices in order to guide educational policies that promote more efficient student learning.

References

Algan, Y., P. Cahuc, and A. Shleifer (2013): "Teaching Practices and Social Capital," *American Economic Journal: Applied Economics*, 5, 189Ű210.

Bietenbeck, J. C. (2014): "Teaching Practices and Cognitive Skills," *Labour Economics*, in press.

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.

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

De Witte, K. and C. Van Klaveren (2014): "How are teachers teaching? A nonparametric approach," *Education Economics*, 22, 3–23.

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

Dobbie, W. and R. G. Fryer (2013): "Getting Beneath the Veil of Effective Schools: Evidence from New York City," *American Economic Journal: Applied Economics*, 5, 28–60.

Fryer, R. G. (2014): "Injecting Charter School Best Practices into Traditional Public Schools: Evidence from field Experiments," *The Quarterly Journal of Economics*, 129, 1355–1407.

Goe, L., C. Bell, and O. Little (2008): "Approaches to Evaluating Teacher Effectiveness: a Research Synthesis," *National Comprehensive Center for Teacher Quality.*

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

Hanushek, E. A. (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 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.

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

Jackson, C. K. (2013): "Non-Cognitive Ability, Test Scores, and Teacher Quality: Evidence from 9th Grade Teachers in North Carolina," NBER Working Paper 18624.

Johnston, D. W., C. Propper, S. E. Pudney, and M. A. Shields (2010): "Is There an Income Gradient in Child Health? It Depends Whom You Ask," IZA Discussion Paper No. 4830.

Lavy, V. (2016): "What Makes an Effective Teacher? Quasi-Experimental Evidence," *CESifo Economic Studies: Oxford Journals*, doi: 10.1093/cesifo/ifv001, 88–125.

OECD (2009): "PISA Data Analysis Manual. SPSS Second Edition," OECD Publications.

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.

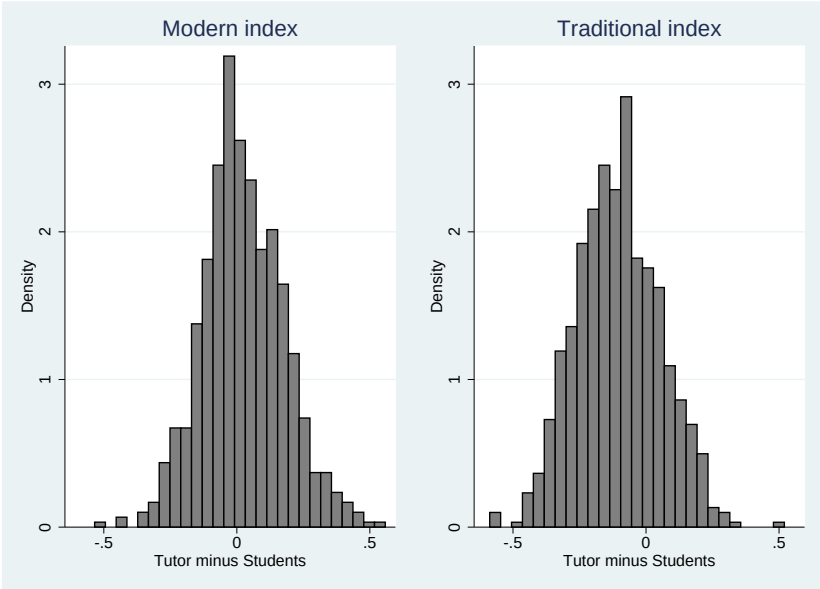
Schwerdt, G. and A. C. Wuppermann (2011): "Is traditional teaching really all that bad? A within-student between-subject approach," *Economics of Education Review*, 30, 365–379.

Van Klaveren, C. (2011): "Lecturing Style Teaching and Student Performance," *Economics of Education Review*, 30, 729–739.

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

Figures

Figure 1: Difference in the teaching indexes between tutor and students



Tables

Table 1: Tutors' characteristics

	Mean	Std. Dev.	Classes†
Female	0.75	0.44	736
Experience (years):			
Less than 5	0.10	0.30	736
5 - 9	0.10	0.30	736
10 - 14	0.07	0.25	736
15 - 19	0.09	0.29	736
20 - 24	0.10	0.30	736
25 - 29	0.15	0.36	736
30 or more	0.39	0.49	736
5-years degree or more	0.17	0.38	736
Instruction:			
Reading and Math	0.88	0.32	736
Reading	0.05	0.21	736
Math	0.07	0.25	736
Person asking for a meeting:			
Parents	0.22	0.41	734
Teacher	0.33	0.47	734
Number of meetings with students' parents	3.04	0.97	731
Teacher at 3rd and 4th grades	0.74	0.44	729
Class size	22.53	3.54	736
Schools			368

†The number of tutors is equal to the number of classes since there is one different tutor per class.

Table 2: Students' characteristics

	Mean	Std. Dev.	Students
Female	0.50	0.50	12113
Repeater	0.06	0.23	12113
Born in 4th quarter	0.32	0.47	12113
Living in single-parent household	0.07	0.26	12113
Living with siblings	0.85	0.36	12113
Country of origin:			
Spain	0.93	0.25	12113
Western Europe	0.00	0.05	12113
Non-Western Europe	0.02	0.12	12113
Morocco	0.01	0.07	12113
Latin America	0.04	0.20	12113
Asia	0.00	0.04	12113
Other	0.00	0.06	12113
Age at starting school:			
2 years old or less	0.60	0.49	12113
3 years old	0.36	0.48	12113
4 years old	0.03	0.16	12113
5 years old	0.01	0.10	12113
6 years old	0.01	0.07	12113
Mother's education:			
Primary or less	0.11	0.32	12113
Compulsory	0.24	0.43	12113
High School	0.14	0.35	12113
Vocational training	0.20	0.40	12113
University	0.31	0.46	12113
Father's education:			
Primary or less	0.13	0.34	12113
Compulsory	0.26	0.44	12113
High School	0.15	0.36	12113
Vocational training	0.20	0.40	12113
University	0.26	0.44	12113
Mother's labor status:			
Self-employed	0.13	0.34	12113
Employee	0.51	0.50	12113
Unemployed	0.10	0.30	12113
Inactive	0.26	0.44	12113
Father's labor status:			
Self-employed	0.26	0.44	12113
Employee	0.65	0.48	12113
Unemployed	0.07	0.26	12113
Inactive	0.02	0.14	12113
Help with homework:			
Private tutor	0.09	0.29	11935
Family	0.61	0.49	11935
Schools			368

Distributions obtained with students weights.

Table 3: Average test scores

	Math	Reading
Full sample	0.14	0.14
	(1.00)	(0.98)
Male students	0.21	0.09
	(1.03)	(0.97)
Female students	0.06	0.19
	(0.96)	(0.98)
Gap (male-female)	0.15	-0.10
Students in public schools	0.04	0.04
	(1.00)	(0.99)
Students in private schools	0.30	0.30
	(0.98)	(0.93)
Gap (public-private)	-0.26	-0.26
Students	12113	
Classes	736	
Schools	368	

Standard deviation in parenthesis. Test scores are standardised with mean 0 and standard deviation 1 in the initial sample.

Table 4: Matched teacher questionnaire items

Traditional Teaching Practices	Modern Teaching Practices
Item (a): Most of the time I teach by telling	Item (b): Students present works or topics to classmates
Item (f): Students work on exercises and activities proposed 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?". Possible answers are "Never or almost never", "Sometimes", "Almost always", and "Always".

Table 5: Correlation across teaching practices reported by the tutor

		Modern items				**Traditional items**			
		(b)	(e)	(h)	(d)	(a)	(f)	(g)	(c)
Modern items	(b)	1.00							
	(e)	0.25***	1.00						
		(0.00)							
	(h)	0.26***	0.26***	1.00					
		(0.00)	(0.00)						
	(d)	0.07	0.18***	0.12***	1.00				
		(0.07)	(0.00)	(0.00)					
Traditional items	(a)	0.09**	-0.04	-0.02	0.09**	1.00			
		(0.01)	(0.24)	(0.60)	(0.01)				
	(f)	0.08	0.08**	0.06	0.25***	0.13***	1.00		
		(0.04)	(0.03)	(0.11)	(0.00)	(0.00)			
	(g)	-0.05	-0.09**	-0.13***	0.09**	0.25***	0.30***	1.00	
		(0.20)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)		
	(c)	0.11***	0.19***	0.07	0.47***	0.19***	0.12***	0.09**	1.00
		(0.00)	(0.00)	(0.06)	(0.00)	(0.00)	(0.00)	(0.02)	

Sample: 368 schools, 736 classrooms, 12113 students. Standard deviation in parenthesis. ** $p < 0.05$, *** $p < 0.01$. (b): "Students present works or topics to classmates"; (e): "I promote discussions"; (h): "Students work in small groups"; (d): "While I teach, students ask me doubts"; (a): "Most of the time I teach by telling"; (f): "Students work on exercises and activities proposed by me"; (g): "Students work individually"; (c): "While I teach, I ask students questions about the lesson".

Table 6: Descriptive statistics of the teaching indexes

	Mean	Std. Dev.	Min.	Max.	Classes
Tutor's answers:					
Modern teaching	0.42	0.14	0.00	0.89	736
Traditional teaching	0.66	0.15	0.11	1.00	736
Students' answers:					
Modern teaching	0.40	0.10	0.13	1.00	736
Traditional teaching	0.76	0.08	0.37	0.95	736

Table 7: Correlation between traditional and modern teaching indexes

		Tutor's answers		Students' answers	
		Mod.	Trad.	Mod.	Trad.
Tutor's answers	Mod.	1.00			
	Trad.	-0.00	1.00		
		(0.89)			
Students' answers	Mod.	0.25***	-0.07**	1.00	
		(0.00)	(0.05)		
	Trad.	-0.09**	0.20***	-0.00	1.00
		(0.02)	(0.00)	(0.90)	

Sample: 368 schools, 736 classrooms, 12113 students. Standard deviation in parenthesis. ** $p < 0.05$, *** $p < 0.01$.

Table 8: Difference in the teaching indexes between tutor and students

	Modern teaching	Traditional teaching
Mean	0.02	-0.11
Standard deviation	0.15	0.16
Median	0.01	-0.11
Skewness	0.15	0.12
Minimum	-0.53	-0.59
Maximum	0.56	0.52
Classes	736	736

Table 9: Variance of teaching indexes

	Tutor's answers		Students' answers†	
	Modern	Traditional	Modern	Traditional
Overall	0.020	0.023	0.010	0.006
Between schools	0.014	0.015	0.007	0.004
Within schools	0.006	0.008	0.003	0.002
% within schools	31.02	33.75	32.88	31.16

Sample: 368 schools, 736 classrooms, 12113 students. †Class average.

Table 10: Within-school sorting: effect of class characteristics (I)

	Dependent variable: Teacher characteristic										
		Years of experience						5-years	Instruction		Tutor
	Female	5-9	10-14	15-19	20-24	25-29	≥30	degree	Math-Reading	Reading	3rd-4th grades
Mother's education:											
Compulsory	0.13	0.13	0.24	0.05	-0.20	-0.42	0.25	0.18	-0.32	0.13	0.20
	(0.35)	(0.33)	(0.31)	(0.33)	(0.23)	(0.31)	(0.45)	(0.27)	(0.28)	(0.25)	(0.38)
High School	0.19	-0.13	0.48*	0.02	-0.06	-0.45	0.22	0.38	-0.16	0.06	0.05
	(0.45)	(0.37)	(0.27)	(0.37)	(0.35)	(0.43)	(0.52)	(0.38)	(0.26)	(0.24)	(0.47)
Vocational training	-0.02	-0.09	0.36	0.00	-0.27	-0.23	0.12	0.09	-0.24	0.03	-0.22
	(0.43)	(0.39)	(0.30)	(0.33)	(0.37)	(0.43)	(0.55)	(0.38)	(0.31)	(0.23)	(0.51)
University	0.09	-0.07	0.38	0.09	-0.28	-0.48	0.17	0.24	-0.31	0.06	-0.35
	(0.45)	(0.37)	(0.31)	(0.31)	(0.34)	(0.39)	(0.56)	(0.42)	(0.31)	(0.28)	(0.49)
Father's education:											
Compulsory	-0.24	-0.01	0.01	0.00	0.25	0.22	-0.16	-0.12	0.56**	-0.30	0.12
	(0.47)	(0.32)	(0.34)	(0.34)	(0.36)	(0.32)	(0.55)	(0.33)	(0.27)	(0.25)	(0.46)
High School	-0.48	-0.01	-0.40	0.18	0.44	0.48	-0.10	-0.37	0.06	0.03	0.30
	(0.50)	(0.35)	(0.28)	(0.28)	(0.38)	(0.40)	(0.55)	(0.43)	(0.26)	(0.26)	(0.43)
Vocational training	-0.39	0.07	-0.23	0.25	0.26	0.57	-0.62	-0.67*	0.23	-0.14	0.12
	(0.47)	(0.34)	(0.31)	(0.31)	(0.39)	(0.39)	(0.54)	(0.39)	(0.29)	(0.28)	(0.46)
University	-0.22	0.27	-0.17	-0.04	0.36	0.39	-0.49	-0.31	0.19	-0.19	0.25
Mother's labor status:											
Employee	0.44	-0.06	-0.12	0.07	0.03	0.20	-0.44	-0.08	-0.12	0.00	-0.21
	(0.30)	(0.24)	(0.21)	(0.24)	(0.22)	(0.28)	(0.35)	(0.29)	(0.16)	(0.14)	(0.30)
Unemployed	0.08	-0.19	0.06	-0.13	0.03	-0.03	0.09	0.18	-0.44*	0.30	-0.26
	(0.41)	(0.37)	(0.32)	(0.30)	(0.32)	(0.34)	(0.51)	(0.42)	(0.27)	(0.20)	(0.43)
Inactive	0.11	-0.01	-0.07	0.07	-0.02	0.27	-0.36	-0.04	-0.08	0.04	-0.11
	(0.38)	(0.26)	(0.22)	(0.30)	(0.29)	(0.35)	(0.46)	(0.33)	(0.21)	(0.17)	(0.39)

Each column is a separate regression, including school fixed effects. Reference outcomes: primary education, self-employed, < 10% non-Spanish. Standard errors clustered at the school level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. (*Continued on next page*)

Table 10: (continued)

	Dependent variable: Teacher characteristic										
		Years of experience						5-years	Instruction		Tutor
	Female	5-9	10-14	15-19	20-24	25-29	≥30	degree	Math-Reading	Reading	3rd-4th grades
Father's labor status:											
Employee	0.07	0.05	0.01	0.16	-0.19	-0.00	0.02	-0.28	0.12	0.03	0.09
	(0.29)	(0.20)	(0.17)	(0.20)	(0.20)	(0.23)	(0.33)	(0.28)	(0.18)	(0.13)	(0.24)
Unemployed	-0.54	0.09	0.04	0.58*	0.20	-0.03	-0.83	-0.83*	0.04	-0.03	0.06
	(0.46)	(0.39)	(0.34)	(0.33)	(0.34)	(0.41)	(0.54)	(0.44)	(0.25)	(0.16)	(0.49)
Inactive	-0.92	-0.74	0.30	-0.03	-0.47	0.77	0.29	0.33	0.64	-0.63	0.50
	(0.89)	(0.56)	(0.48)	(0.51)	(0.55)	(0.75)	(1.14)	(0.60)	(0.45)	(0.42)	(0.80)
% non-Spanish											
10-20%	0.02	0.07	0.02	-0.07	-0.07	-0.02	0.04	0.14*	-0.01	-0.00	0.05
	(0.09)	(0.07)	(0.05)	(0.06)	(0.06)	(0.09)	(0.11)	(0.07)	(0.05)	(0.03)	(0.09)
More 20%	-0.07	0.05	0.10	-0.04	-0.06	0.01	-0.12	0.24*	-0.05	0.04	-0.02
	(0.12)	(0.11)	(0.08)	(0.08)	(0.12)	(0.12)	(0.17)	(0.13)	(0.10)	(0.06)	(0.14)
% single parent	-0.15	-0.06	0.06	0.01	-0.32	0.13	0.47	-0.17	0.11	0.02	0.32
	(0.42)	(0.30)	(0.27)	(0.28)	(0.37)	(0.38)	(0.54)	(0.46)	(0.24)	(0.22)	(0.44)
% siblings	-0.22	0.19	0.02	-0.12	-0.07	-0.01	0.14	-0.20	0.05	-0.13	0.37
	(0.29)	(0.25)	(0.17)	(0.15)	(0.22)	(0.26)	(0.33)	(0.32)	(0.21)	(0.15)	(0.31)
% female	-0.44*	0.11	0.09	0.00	-0.09	-0.07	0.00	0.07	-0.07	-0.04	0.29
	(0.26)	(0.18)	(0.17)	(0.19)	(0.19)	(0.24)	(0.31)	(0.24)	(0.16)	(0.17)	(0.24)
% repeater	0.17	0.34	-0.03	-0.07	0.10	-0.27	-0.21	0.24	0.49*	-0.28	-0.18
	(0.54)	(0.33)	(0.48)	(0.34)	(0.38)	(0.43)	(0.56)	(0.39)	(0.28)	(0.24)	(0.55)
F-test	0.67	0.40	0.50	0.40	0.49	0.47	0.58	0.77	0.64	0.38	0.38
p-value	0.85	0.99	0.97	0.99	0.97	0.98	0.92	0.75	0.88	0.99	0.99
R^2	0.59	0.57	0.57	0.60	0.51	0.55	0.56	0.55	0.75	0.59	0.64
Classes	736	736	736	736	736	736	736	736	736	736	729

Each column is a separate regression, including school fixed effects. Reference outcomes: primary education, self-employed, < 10% non-Spanish. Standard errors clustered at the school level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. F-test: joint significance of the class characteristics.

Table 11: Within-school sorting: effect of class characteristics (II)

	Teaching index (Tutor's answers)		Teaching index (Students' answers)	
	Traditional	Modern	Traditional	Modern
Mother's education:				
Compulsory	-0.04	-0.10	0.01	-0.06
	(0.13)	(0.12)	(0.09)	(0.12)
High School	0.02	-0.09	-0.03	-0.11
	(0.17)	(0.14)	(0.10)	(0.14)
Vocational training	-0.15	-0.07	-0.04	-0.06
	(0.17)	(0.14)	(0.10)	(0.14)
University	0.06	-0.13	-0.02	-0.06
	(0.16)	(0.14)	(0.11)	(0.15)
Father's education:				
Compulsory	0.09	0.04	0.00	0.03
	(0.14)	(0.13)	(0.09)	(0.11)
High School	0.13	-0.01	0.06	0.00
	(0.15)	(0.13)	(0.08)	(0.11)
Vocational training	0.10	0.04	0.05	0.00
	(0.14)	(0.13)	(0.08)	(0.11)
University	-0.03	0.07	0.03	-0.01
	(0.14)	(0.13)	(0.09)	(0.11)
Mother's labor status:				
Employee	0.06	0.00	-0.05	-0.03
	(0.09)	(0.09)	(0.06)	(0.07)
Unemployed	0.14	-0.08	-0.05	-0.01
	(0.13)	(0.11)	(0.09)	(0.10)
Inactive	0.12	-0.05	0.01	-0.04
	(0.13)	(0.12)	(0.07)	(0.08)
Father's labor status:				
Employee	0.02	-0.14**	-0.01	0.02
	(0.08)	(0.07)	(0.05)	(0.07)
Unemployed	-0.07	-0.26**	-0.04	-0.08
	(0.14)	(0.12)	(0.08)	(0.11)
Inactive	0.12	-0.16	-0.18	-0.01
	(0.24)	(0.20)	(0.12)	(0.17)
% non-Spanish				
10-20%	0.00	0.01	0.01	0.03
	(0.03)	(0.03)	(0.01)	(0.02)
More 20%	-0.01	0.04	0.01	0.03
	(0.04)	(0.04)	(0.02)	(0.03)
% single parent	-0.01	-0.17	0.00	-0.06
	(0.13)	(0.13)	(0.07)	(0.12)
% siblings	0.08	0.00	0.02	0.08
	(0.10)	(0.08)	(0.05)	(0.07)
% female	-0.02	0.03	0.04	0.01
	(0.08)	(0.08)	(0.05)	(0.06)
% repeater	0.08	0.01	-0.02	0.14
	(0.14)	(0.14)	(0.09)	(0.12)
F-test	0.43	0.69	0.49	0.47
p-value	0.99	0.83	0.97	0.98
School fixed effects	Yes	Yes	Yes	Yes
R^2	0.68	0.71	0.70	0.69
Classes	736	736	736	736

Reference outcomes: primary education, self-employed, $< 10\%$ non-Spanish. Standard errors clustered at the school level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. F-test: joint significance of the class characteristics.

Table 12: Within-school sorting: effect of class size

	Traditional teaching		Modern teaching	
Class size	-0.01	-0.03	-0.02	-0.00
	(0.02)	(0.03)	(0.01)	(0.03)
(Class size)^2	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
F-test	0.38	0.59	1.27	0.25
p-value	0.69	0.56	0.28	0.78
School fixed effects	No	Yes	No	Yes
R^2	0.00	0.66	0.01	0.69
Classes	736	736	736	736

Teaching practices reported by the tutor. Standard errors clustered at the school level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. F-test: joint significance of class size and (class size)^2.

Table 13: Within-school selection of teaching style: Effect of teacher characteristics

	Dependent variable: Teaching index			
	Tutor's answers		Students' answers	
	Traditional	Modern	Traditional	Modern
Female	-0.02	0.01	-0.02*	0.00
	(0.03)	(0.02)	(0.01)	(0.02)
Years of experience (ref: < 5):				
5 - 9	0.03	0.01	-0.01	0.01
	(0.04)	(0.04)	(0.02)	(0.03)
10 - 14	0.04	0.00	-0.02	-0.03
	(0.05)	(0.05)	(0.03)	(0.04)
15 - 19	0.03	0.01	-0.01	-0.02
	(0.05)	(0.05)	(0.02)	(0.03)
20 - 24	0.02	-0.00	-0.00	-0.02
	(0.05)	(0.05)	(0.02)	(0.03)
25 - 29	0.01	0.01	-0.02	-0.03
	(0.04)	(0.04)	(0.02)	(0.03)
30 or more	0.05	0.01	-0.00	-0.04
	(0.04)	(0.04)	(0.02)	(0.02)
5-years degree or more	-0.03	-0.00	-0.00	0.01
	(0.03)	(0.03)	(0.01)	(0.02)
Instruction (ref: Math):				
Reading and Math	-0.02	-0.01	-0.01	0.03
	(0.06)	(0.05)	(0.03)	(0.03)
Reading	-0.00	0.01	0.01	0.04
	(0.08)	(0.07)	(0.03)	(0.04)
Person asking for a meeting:				
Parents	0.01	-0.01	-0.01	0.01
	(0.04)	(0.03)	(0.02)	(0.02)
Teacher	-0.01	0.00	-0.00	0.02
	(0.03)	(0.03)	(0.01)	(0.02)
# of meetings with parents	0.02	-0.00	-0.01	0.01
	(0.02)	(0.02)	(0.01)	(0.01)
Teacher at 3rd and 4th grades	-0.01	-0.01	0.02	0.02
	(0.03)	(0.03)	(0.02)	(0.02)
Class size	-0.00	0.00	0.00	0.00
	(0.01)	(0.01)	(0.00)	(0.01)
Constant	0.55***	0.51***	0.74***	0.30**
	(0.18)	(0.16)	(0.10)	(0.14)
F-test	0.46	0.06	0.59	0.71
p-value	0.96	1.00	0.88	0.78
School fixed effects	Yes	Yes	Yes	Yes
R^2	0.69	0.69	0.71	0.69
Classes	724	724	724	724

Standard errors clustered at the school level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
F-test: joint significance of teacher characteristics and class size.

Table 14: Estimation results

	Tutor's answers			Students' answers		
	(1)	(2)	(3)	(4)	(5)	(6)
Traditional teaching	0.00	0.04	0.00	-0.19	-0.21	-0.26**
	(0.10)	(0.10)	(0.10)	(0.13)	(0.13)	(0.13)
Modern teaching	0.08	0.08	0.14	0.38***	0.47***	0.34**
	(0.10)	(0.10)	(0.09)	(0.14)	(0.14)	(0.13)
Math dummy	0.00	-0.00	-0.00	-0.00	-0.00	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Class size		0.01*	0.01		0.01	0.00
		(0.01)	(0.01)		(0.01)	(0.01)
Teacher variables:						
Female		0.03	0.01		0.02	0.01
		(0.02)	(0.02)		(0.03)	(0.03)
Years of experience (ref: < 5):						
5 - 9		0.02	0.03		0.01	0.02
		(0.04)	(0.04)		(0.05)	(0.04)
10 - 14		0.07	0.07		0.04	0.07
		(0.06)	(0.05)		(0.06)	(0.06)
15 - 19		0.07	0.09*		0.07	0.09*
		(0.05)	(0.05)		(0.05)	(0.05)
20 - 24		0.01	0.01		-0.00	0.02
		(0.05)	(0.05)		(0.05)	(0.05)
25 - 29		0.06	0.06		0.04	0.05
		(0.05)	(0.04)		(0.05)	(0.05)
30 or more		0.07*	0.08*		0.08*	0.10**
		(0.04)	(0.04)		(0.04)	(0.04)
5-years degree or more		-0.07***	-0.07***		-0.08***	-0.07***
		(0.03)	(0.03)		(0.03)	(0.03)
Instruction (ref: Math):						
Reading and Math		-0.10*	-0.04		-0.08	-0.03
		(0.06)	(0.05)		(0.05)	(0.05)
Reading		-0.15**	-0.13**		-0.15**	-0.12*
		(0.06)	(0.06)		(0.06)	(0.06)
Person asking for a meeting:						
Parents		-0.01	0.01		-0.03	-0.01
		(0.03)	(0.03)		(0.04)	(0.03)
Teacher		-0.02	0.01		-0.03	0.00
		(0.03)	(0.03)		(0.04)	(0.04)
# of meetings with parents		-0.02	-0.02		-0.02	-0.02
		(0.02)	(0.02)		(0.02)	(0.02)
Teacher at 3rd and 4th grades		0.02	0.05*		0.02	0.03
		(0.03)	(0.02)		(0.03)	(0.03)
Constant	0.08	-0.14	-0.27	0.16*	0.06	0.04
	(0.08)	(0.20)	(0.20)	(0.09)	(0.20)	(0.20)
Student characteristics	No	No	Yes	No	No	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	24226	23844	23492	22086	21734	21524
R^2	0.16	0.16	0.23	0.15	0.15	0.22

Dependent variable: Student test scores in Math and reading. Standard errors clustered at the class level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Student's characteristics: female, country of origin, repeater, mother and father's education, mother and father's labor status, single-parent household, siblings, born in 4th quarter, age at starting school, private tutor/family helps with homework.

Table 15: Heterogeneous effect by subject

	Tutor's answers		Students' answers	
	Math	Reading	Math	Reading
Traditional teaching	0.08	-0.07	-0.26	-0.26*
	(0.12)	(0.10)	(0.18)	(0.14)
Modern teaching	-0.03	0.30***	0.27*	0.41**
	(0.13)	(0.10)	(0.15)	(0.16)
Class size	0.01	0.01	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)
Teacher variables:				
Female	-0.01	0.04	-0.01	0.03
	(0.03)	(0.02)	(0.04)	(0.03)
Years of experience (ref: < 5):				
5 - 9	-0.03	0.10*	-0.05	0.10*
	(0.05)	(0.05)	(0.06)	(0.06)
10 - 14	0.01	0.12*	0.02	0.12
	(0.07)	(0.07)	(0.07)	(0.08)
15 - 19	0.12**	0.05	0.10*	0.08
	(0.05)	(0.06)	(0.06)	(0.07)
20 - 24	0.04	-0.02	0.03	0.00
	(0.05)	(0.06)	(0.06)	(0.07)
25 - 29	0.06	0.07	0.03	0.08
	(0.06)	(0.06)	(0.06)	(0.06)
30 or more	0.07	0.08	0.07	0.12**
	(0.05)	(0.05)	(0.06)	(0.06)
5-years degree or more	-0.06**	-0.08**	-0.07**	-0.08**
	(0.03)	(0.03)	(0.03)	(0.03)
Instruction (ref: Math):				
Reading and Math	-0.07	-0.02	-0.07	0.00
	(0.06)	(0.07)	(0.07)	(0.07)
Reading	-0.17**	-0.08	-0.19**	-0.05
	(0.07)	(0.08)	(0.08)	(0.08)
Person asking for a meeting:				
Parents	0.01	0.00	-0.01	-0.01
	(0.04)	(0.03)	(0.04)	(0.04)
Teacher	-0.00	0.03	-0.01	0.02
	(0.04)	(0.04)	(0.04)	(0.04)
# of meetings with parents	-0.03	-0.00	-0.03	-0.00
	(0.03)	(0.02)	(0.03)	(0.03)
Teacher at 3rd and 4th grades	0.05*	0.04	0.04	0.01
	(0.03)	(0.04)	(0.03)	(0.04)
Constant	-0.07	-0.47*	0.28	-0.21
	(0.26)	(0.25)	(0.30)	(0.28)
Student characteristics	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes
Observations	11746	11746	10762	10762
R^2	0.24	0.26	0.23	0.24

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

Table 16: Heterogeneous effects: top and bottom performers

	A. Tutor's answers			
	Math		Reading	
	Top	Bottom	Top	Bottom
Traditional teaching	0.18	-0.03	-0.01	-0.11
	(0.18)	(0.12)	(0.11)	(0.13)
Modern teaching	0.08	-0.22	0.16	0.02
	(0.16)	(0.26)	(0.13)	(0.11)
Observations	2982	2948	3013	2877
R^2	0.18	0.17	0.17	0.21

	B. Students' answers			
	Math		Reading	
	Top	Bottom	Top	Bottom
Traditional teaching	-0.17	-0.08	0.03	-0.07
	(0.20)	(0.20)	(0.22)	(0.26)
Modern teaching	-0.17	-0.12	0.21	0.31*
	(0.24)	(0.23)	(0.20)	(0.19)
Observations	2867	2524	2875	2466
R^2	0.19	0.18	0.18	0.22
Student characteristics	Yes	Yes	Yes	Yes
Teacher characteristics	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes

Each column in each panel represents a separate regression. Top performers: Students above the 75 percentile of the distribution of the first plausible value. Bottom performers: Students below the 25 percentile of the distribution of the first plausible value. Standard errors clustered at the class level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 17: Heterogeneous effects by gender of students, type of school and type of tutor

	Math and Reading		**Math**		**Reading**	
	A. Gender of students					
	Boys	Girls	Boys	Girls	Boys	Girls
Tutor's answers:	[11844]	[11648]	[5922]	[5824]	[5922]	[5824]
Traditional teaching	0.04	-0.06	0.06	0.05	0.02	-0.16
	(0.11)	(0.13)	(0.14)	(0.16)	(0.12)	(0.14)
Modern teaching	0.11	0.17	-0.07	0.05	0.30**	0.29**
	(0.12)	(0.12)	(0.17)	(0.15)	(0.14)	(0.13)
Math dummy	0.12***	-0.13***	-	-	-	-
	(0.02)	(0.02)	-	-	-	-
Students' answers:	[10860]	[10664]	[5430]	[5332]	[5430]	[5332]
Traditional teaching	0.02	-0.56***	0.01	-0.57**	0.04	-0.55**
	(0.18)	(0.21)	(0.22)	(0.27)	(0.19)	(0.24)
Modern teaching	0.13	0.44**	0.12	0.29	0.14	0.60**
	(0.18)	(0.18)	(0.22)	(0.21)	(0.19)	(0.24)
Math dummy	0.12***	-0.13***	-	-	-	-
	(0.02)	(0.02)	-	-	-	-
	B. Type of school					
	Public	Private	Public	Private	Public	Private
Tutor's answers:	[15000]	[8492]	[7500]	[4246]	[7500]	[4246]
Traditional teaching	-0.07	0.29	-0.03	0.45	-0.12	0.14
	(0.10)	(0.21)	(0.11)	(0.28)	(0.12)	(0.18)
Modern teaching	0.13	0.09	-0.04	-0.02	0.29*	0.20
	(0.13)	(0.15)	(0.20)	(0.18)	(0.16)	(0.15)
Math dummy	0.00	-0.01	-	-	-	-
	(0.01)	(0.02)	-	-	-	-
Students' answers:	[13590]	[7934]	[6795]	[3967]	[6795]	[3967]
Traditional teaching	-0.29**	-0.01	-0.26	-0.15	-0.32*	0.14
	(0.15)	(0.25)	(0.20)	(0.30)	(0.18)	(0.25)
Modern teaching	0.40**	0.18	0.42**	-0.03	0.37*	0.38
	(0.16)	(0.26)	(0.19)	(0.32)	(0.21)	(0.25)
Math dummy	-0.00	-0.01	-	-	-	-
	(0.02)	(0.02)	-	-	-	-
	C. Type of tutor (in 3rd and 4th grades; only in 4th grade)					
	4th grade	3rd-4th grades	4th grade	3rd-4th grades	4th grade	3rd-4th grades
Tutor's answers:	[5666]	[17826]	[2833]	[8913]	[2833]	[8913]
Traditional teaching	0.06	0.07	-0.05	0.13	0.17	0.01
	(0.27)	(0.11)	(0.35)	(0.14)	(0.25)	(0.13)
Modern teaching	0.47	0.24*	0.39	0.10	0.55	0.37***
	(0.43)	(0.13)	(0.49)	(0.17)	(0.47)	(0.14)
Math dummy	-0.05	0.01	-	-	-	-
	(0.03)	(0.02)	-	-	-	-
Students' answers:	[5222]	[16302]	[2611]	[8151]	[2611]	[8151]
Traditional teaching	-0.98*	-0.28	-0.58	-0.37	-1.38**	-0.19
	(0.52)	(0.19)	(0.68)	(0.25)	(0.68)	(0.20)
Modern teaching	1.73***	0.44**	2.02***	0.40**	1.44*	0.48*
	(0.67)	(0.19)	(0.75)	(0.19)	(0.74)	(0.26)
Math dummy	-0.05*	0.01	-	-	-	-
	(0.03)	(0.01)	-	-	-	-

Observations in brackets. Standard errors clustered at the class level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All regressions control for teacher and student characteristics, and school fixed effects.

Table 18: Difference between the student's modern and traditional indexes

Math score	-0.02
	(0.01)
Reading score	0.03**
	(0.02)
Math score (class average)†	-0.02***
	(0.00)
Reading score (class average)†	-0.03***
	(0.00)
Public school	0.13***
	(0.05)
Selected student characteristics††:	
Female	-0.03***
	(0.01)
Repeater	0.06***
	(0.01)
Born in 4th quarter	-0.01**
	(0.01)
Age at starting school (ref: ≤2 years old):	
3 years old	-0.00
	(0.01)
4 years old	-0.02
	(0.02)
5 years old	0.06**
	(0.03)
6 years old	0.00
	(0.03)
Mother's education (ref: Primary or less)	
Compulsory	-0.03**
	(0.01)
High School	-0.02
	(0.01)
Vocational training	-0.02
	(0.01)
University	-0.04***
	(0.01)
Help with homework:	
Private tutor	0.04***
	(0.01)

Standard errors clustered at the class level in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$,*** $p < 0.01$. †Excluding student's own value. ††The following characteristics are also included: siblings, single-parent household, country of origin, parents' labor status, father's education, experience, 5-years degree or more, person asking for meeting, meetings with parents, tutor in 3rd and 4th grades. For the sake of brevity we report characteristics with significant coefficients (the rest are available upon request). (*Continued on next page*)

Table 18: (continued)

Family	0.01**
	(0.01)
Selected tutor characteristics††:	
Female	0.02***
	(0.01)
Instruction (ref: Math):	
Reading and Math	0.04**
	(0.02)
Reading	0.03*
	(0.02)
Constant	-0.41***
	(0.10)
School fixed effects	Yes
Class-average of student characteristics†	Yes
Observations	10762
R^2	0.19

Standard errors clustered at the class level in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$,*** $p < 0.01$. †Excluding student's own value. ††The following characteristics are also included: siblings, single-parent household, country of origin, parents' labor status, father's education, experience, 5-years degree or more, person asking for meeting, meetings with parents, tutor in 3rd and 4th grades. For the sake of brevity we report characteristics with significant coefficients (the rest are available upon request).

Table 19: Sensitivity to include each index separately

	Math-Reading			**Math**			**Reading**		
	Baseline	(1)	(2)	Baseline	(3)	(4)	Baseline	(5)	(6)
A. Tutor's answers:									
Traditional teaching	0.00	0.01		0.08	0.08		-0.07	-0.06	
	(0.10)	(0.10)		(0.12)	(0.12)		(0.10)	(0.10)	
Modern teaching	0.14		0.14	-0.03		-0.02	0.30***		0.29***
	(0.09)		(0.09)	(0.13)		(0.13)	(0.10)		(0.10)
Observations	23492	23492	23492	11746	11746	11746	11746	11746	11746
B. Students' answers:									
Traditional teaching	-0.26**	-0.18		-0.26	-0.22		-0.26*	-0.15	
	(0.13)	(0.12)		(0.18)	(0.17)		(0.14)	(0.12)	
Modern teaching	0.34**		0.27**	0.27*		0.19	0.41**		0.36**
	(0.13)		(0.12)	(0.15)		(0.14)	(0.16)		(0.15)
Observations	21524	22166	22022	10762	11083	11011	10762	11083	11011

Baseline columns report results from Tables 14 and 15. Each column in each Panel A and B represents a separate regression. All regressions control for student and teacher characteristics, class size and school fixed effects. Standard errors clustered at the class level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 20: Sensitivity to include the teaching practices individually

	Math and Reading		**Math**		**Reading**	
	Tutor	Students	Tutor	Students	Tutor	Students
	A. Baseline estimates					
Traditional teaching	0.00	-0.26**	0.08	-0.07	-0.26	-0.26*
	(0.10)	(0.13)	(0.12)	(0.10)	(0.18)	(0.14)
Modern teaching	0.14	0.34**	-0.03	0.30***	0.27*	0.41**
	(0.09)	(0.13)	(0.13)	(0.10)	(0.15)	(0.16)
Observations	23492	21524	11746	10762	11746	10762
	B. Without aggregating individual practices					
Traditional practices						
Teach by telling	0.01	0.12	0.05	0.09	-0.02	0.14
	(0.06)	(0.12)	(0.07)	(0.15)	(0.06)	(0.14)
Exercises proposed by teacher	-0.14**	-0.10	-0.14	-0.17	-0.15**	-0.03
	(0.07)	(0.16)	(0.09)	(0.19)	(0.07)	(0.17)
Students work individually	0.13**	-0.18	0.16**	-0.06	0.11	-0.31**
	(0.06)	(0.14)	(0.08)	(0.19)	(0.08)	(0.13)
Modern practices						
Student present works/topics	-0.03	0.21**	-0.06	0.15	0.01	0.27**
	(0.08)	(0.10)	(0.11)	(0.13)	(0.09)	(0.12)
Teacher promotes discussions	0.09**	0.02	0.06	-0.03	0.12**	0.07
	(0.04)	(0.10)	(0.05)	(0.13)	(0.05)	(0.10)
Students work in small groups	0.07	0.09	-0.04	0.14	0.18**	0.04
	(0.08)	(0.10)	(0.09)	(0.14)	(0.09)	(0.12)
Math	-0.00	-0.01				
	(0.01)	(0.01)				
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Tutor characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Student characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	23492	21524	11746	10762	11746	10762
R^2	0.23	0.22	0.24	0.23	0.26	0.24

Each column in each Panel A and B represents a separate regression. Standard errors clustered at the class level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 21: Sensitivity to alternative measures of the teaching style and to additional controls

	Math-Reading		**Math**		**Reading**	
	Tutor	Students	Tutor	Students	Tutor	Students
	A. Baseline estimates					
Traditional teaching	0.00	-0.26**	0.08	-0.07	-0.26	-0.26*
	(0.10)	(0.13)	(0.12)	(0.10)	(0.18)	(0.14)
Modern teaching	0.14	0.34**	-0.03	0.30***	0.27*	0.41**
	(0.09)	(0.13)	(0.13)	(0.10)	(0.15)	(0.16)
Observations	23492	21524	11746	10762	11746	10762
	B. Including items (c) and (d) in the teaching indexes†					
B1. (c) traditional; (d) modern						
Traditional teaching	0.07	-0.31**	0.15	-0.23	-0.01	-0.38**
	(0.12)	(0.16)	(0.14)	(0.22)	(0.13)	(0.17)
Modern teaching	0.20**	0.30**	0.05	0.22	0.35***	0.39**
	(0.10)	(0.15)	(0.13)	(0.18)	(0.11)	(0.17)
B2. (c) modern; (d) traditional						
Traditional teaching	0.09	-0.40***	0.20	-0.46**	-0.02	-0.34**
	(0.11)	(0.15)	(0.13)	(0.23)	(0.14)	(0.17)
Modern teaching	0.18*	0.38***	-0.00	0.40**	0.36***	0.35**
	(0.10)	(0.14)	(0.14)	(0.17)	(0.11)	(0.17)
Observations	22064	20958	11032	10479	11032	10479
	C. Restricting total time allocated to modern and traditional items					
% of time using modern practices††	0.20	0.39**	-0.07	0.28	0.47***	0.50**
	(0.12)	(0.19)	(0.15)	(0.22)	(0.17)	(0.25)
Observations	23492	21508	11746	10754	11746	10754
	D. Additional controls: class-average of student characteristics†††					
Traditional teaching	0.02	-0.27**	0.08	-0.28	-0.04	-0.26*
	(0.09)	(0.13)	(0.11)	(0.19)	(0.11)	(0.14)
Modern teaching	0.11	0.37***	-0.03	0.31**	0.25**	0.43***
	(0.09)	(0.13)	(0.12)	(0.15)	(0.10)	(0.16)
Observations	23492	21524	11746	10762	11746	10762

Each column in each Panel A, B1, B2, C and D represents a separate regression. All regressions control for student and teacher characteristics, class size and school fixed effects. †Item (c): "While I teach, I ask students questions about the lesson"; item (d): "While I teach, students ask me doubts". ††The remaining percentage corresponds to using traditional practices. †††Excluding the student's own value. Standard errors clustered at the class level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.1: Characteristics of teachers and students in initial and final samples

	Initial	Final
Teachers:		
Female	0.74	0.75
Experience (years):		
Less than 5	0.11	0.10
5 - 9	0.12	0.10
10 - 14	0.08	0.07
15 - 19	0.09	0.09
20 - 24	0.10	0.10
25 - 29	0.16	0.15
30 or more	0.35	0.39
5-years degree or more	0.18	0.17
Instruction:		
Reading and Math	0.84	0.88
Reading	0.05	0.05
Math	0.08	0.07
Person asking for a meeting:		
Parents	0.19	0.22
Teacher	0.37	0.33
# of meetings with parents	3.11	3.04
Teacher at 3rd and 4th grades	0.69	0.74
Students:		
Female	0.49	0.50
Repeater	0.08	0.06
Born in 4th quarter	0.33	0.32
Living in single-parent household	0.10	0.07
Living with siblings	0.83	0.85
Country of origin:		
Spain	0.90	0.93
Western Europe	0.00	0.00
Non-Western Europe	0.02	0.02
Morocco	0.01	0.01
Latin America	0.06	0.04
Asia	0.00	0.00
Other	0.01	0.00
Age at starting school:		
2 years old or less	0.58	0.60
3 years old	0.37	0.36
4 years old	0.03	0.03
5 years old	0.01	0.01
6 years old	0.01	0.01
Mother's education:		
Primary or less	0.13	0.11

Table A.1: (continued)

	Initial	Final
Compulsory	0.26	0.24
High School	0.14	0.14
Vocational training	0.19	0.20
University	0.27	0.31
Father's education:		
Primary or less	0.15	0.13
Compulsory	0.27	0.26
High School	0.15	0.15
Vocational training	0.20	0.20
University	0.23	0.26
Mother's labor status:		
Self-employed	0.14	0.13
Employee	0.50	0.51
Unemployed	0.11	0.10
Inactive	0.25	0.26
Father's labor status:		
Self-employed	0.26	0.26
Employee	0.63	0.65
Unemployed	0.09	0.07
Inactive	0.02	0.02
Help with homework:		
Private tutor	0.10	0.09
Family	0.60	0.61
Students	28708	12113
Classes	1358	736
Schools	900	368
% public schools	69.27	69.29

Table A.2: Distribution of the teaching indexes

	Mean	Std. Dev.	Min.	Max.	Classes
Initial sample					
Modern index (tutor)	0.43	0.14	0.00	0.89	1317
Traditional index (tutor)	0.64	0.15	0.11	1.00	1308
Modern index (students)	0.40	0.09	0.11	1.00	1357
Traditional index (students)	0.75	0.08	0.38	1.00	1357
Final sample					
Modern index (tutor)	0.42	0.14	0.00	0.89	736
Traditional index (tutor)	0.66	0.15	0.11	1.00	736
Modern index (students)	0.40	0.10	0.13	1.00	736
Traditional index (students)	0.76	0.08	0.37	0.95	736

Table A.3: Estimation results of student characteristics

	Tutor's answers			Students' answers		
	M-R	Math	Reading	M-R	Math	Reading
Female	-0.02	-0.14***	0.11***	-0.02	-0.15***	0.10***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Repeater	-0.35***	-0.31***	-0.38***	-0.35***	-0.32***	-0.39***
	(0.03)	(0.04)	(0.05)	(0.04)	(0.05)	(0.05)
Single-parent household	-0.07**	-0.06	-0.08**	-0.08**	-0.06	-0.10**
	(0.03)	(0.04)	(0.04)	(0.03)	(0.05)	(0.04)
Siblings	-0.06***	-0.04	-0.07**	-0.05***	-0.04	-0.07**
	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
Born in 4th quarter	-0.15***	-0.15***	-0.15***	-0.15***	-0.15***	-0.14***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Age at starting school (ref: ≤2 years old):						
3 years old	-0.04**	-0.03	-0.05**	-0.05**	-0.04	-0.05**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
4 years old	-0.09*	-0.07	-0.12*	-0.08	-0.06	-0.11*
	(0.05)	(0.07)	(0.06)	(0.05)	(0.07)	(0.06)
5 years old	-0.20***	-0.12	-0.29***	-0.22***	-0.14	-0.30***
	(0.06)	(0.09)	(0.08)	(0.08)	(0.10)	(0.09)
6 years old	-0.41***	-0.32***	-0.50***	-0.38***	-0.28**	-0.47***
	(0.09)	(0.12)	(0.11)	(0.10)	(0.14)	(0.12)
Country of origin (ref: Spain):						
Western Europe	-0.06	0.20	-0.33	-0.02	0.23	-0.27
	(0.15)	(0.19)	(0.20)	(0.16)	(0.21)	(0.21)
Non-Western Europe	-0.16***	-0.13*	-0.18***	-0.18***	-0.16**	-0.20***
	(0.05)	(0.07)	(0.07)	(0.05)	(0.07)	(0.07)
Morocco	-0.19*	-0.17	-0.21*	-0.10	-0.12	-0.09
	(0.10)	(0.14)	(0.11)	(0.11)	(0.15)	(0.13)
Latin America	-0.21***	-0.22***	-0.20***	-0.22***	-0.22***	-0.22***
	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)
Asia	-0.19	-0.13	-0.24	-0.16	-0.07	-0.26
	(0.18)	(0.25)	(0.23)	(0.18)	(0.26)	(0.25)
Other	-0.29**	-0.23*	-0.35**	-0.32**	-0.26	-0.37**
	(0.11)	(0.14)	(0.15)	(0.13)	(0.16)	(0.17)
Mother's education (ref: Primary or less):						
Compulsory	0.10***	0.10***	0.10***	0.09***	0.08**	0.09**
	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)
High School	0.18***	0.16***	0.20***	0.19***	0.16***	0.22***
	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)
Vocational training	0.15***	0.13***	0.17***	0.16***	0.14***	0.18***
	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)
University	0.33***	0.32***	0.34***	0.32***	0.31***	0.33***

M-R: Math and Reading. Standard errors clustered at the class level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

(*Continued on next page*)

	Tutor's answers			Students' answers		
	M-R	Math	Reading	M-R	Math	Reading
	(0.03)	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)
Mother's labor status (ref: Self-employed):						
Employee	0.06**	0.07**	0.06**	0.06**	0.06*	0.05*
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Unemployed	0.03	0.04	0.02	0.03	0.03	0.02
	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)
Inactive	0.04	0.04	0.03	0.03	0.04	0.01
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Father's education (ref: Primary or less):						
Compulsory	0.08***	0.05	0.11***	0.06**	0.04	0.08**
	(0.03)	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)
High School	0.21***	0.16***	0.25***	0.19***	0.15***	0.22***
	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)
Vocational training	0.17***	0.13***	0.21***	0.15***	0.13***	0.18***
	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)	(0.04)
University	0.30***	0.26***	0.35***	0.28***	0.24***	0.32***
	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)
Father's labor status (ref: Self-employed):						
Employee	0.01	-0.01	0.03	-0.00	-0.02	0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
Unemployed	-0.04	-0.05	-0.03	-0.05	-0.08*	-0.03
	(0.03)	(0.04)	(0.05)	(0.03)	(0.04)	(0.04)
Inactive	-0.04	-0.06	-0.01	-0.05	-0.07	-0.03
	(0.06)	(0.07)	(0.08)	(0.06)	(0.07)	(0.07)
Help with homework:						
Private tutor	-0.42***	-0.40***	-0.44***	-0.42***	-0.40***	-0.43***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)
Family	-0.08***	-0.11***	-0.06***	-0.08***	-0.11***	-0.05***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Constant	-0.27	-0.07	-0.47*	0.04	0.28	-0.21
	(0.20)	(0.26)	(0.25)	(0.20)	(0.30)	(0.28)
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Teacher characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	23492	11746	11746	21524	10762	10762
R^2	0.23	0.24	0.26	0.22	0.23	0.24

M-R: Math and Reading. Standard errors clustered at the class level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.