From Teacher Quality to Teaching Quality: Instructional Productivity and Teaching Practices in the US*

Simon Briole†

Paris School of Economics

November 9, 2020

Abstract

Though teachers are consistently found to play a major role in determining student achievement, little is known about what teachers can do to increase their instructional productivity. This paper develops a new empirical strategy, based on *within-student within math* variations in student test scores, to assess the instructional hourly productivity of math teachers in the US. Building on these estimates, I show that teachers' hourly productivity strongly relates to the use of teaching practices emphasizing student active participation in the lesson (*modern practices*). One weekly hour of math instructional time increases student test scores by 4.4% of a standard deviation on average, but one hour spent with a teacher above the modern practices index median is more than twice as productive as one hour spent with a teacher under this median (+5.9% vs +2.7% standard deviations). A further investigation suggests that the positive effects associated to modern practices are partially mediated by an improvement in student self-confidence and motivation to learn mathematics.

JEL classification: I20; I21; J24

Keywords: teacher quality ; teaching practices; instruction time; TIMSS; test scores; education

---

*I would like to thank Eric Maurin for his support and all his helpful comments. I am also grateful to Elise Huillery, Manon Garrouste, Marc Gurgand, Sandra McNally, Andrew Clark, Daniel Waldenstrőm, Philipp Ketz, Pierre Courtioux, Thomas Breda, François Libois, Kenneth Houngbedji, Marion Davin and all seminar participants at the EEA 2017 conference (Lisbon), the EALE 2017 conference (St Gallen), the JMA 2017 conference (Le Mans), the ADRES 2019 conference (Lyon) as well as participants to workshops and internal seminars at the Paris School of Economics, Pompeu Fabra University, Aix-Marseille School of Economics and University of Montpellier for their valuable comments.

†simon.briole@psemail.eu

Introduction

It is well established that teachers differ a lot in their individual capacity to raise student test scores (Rockoff (2004), Rivkin et al. (2005), Hanushek & Rivkin (2006) and Hanushek & Rivkin (2010)). Furthermore, being taught by a good teacher matters beyond schooling as it positively affects adult outcomes such as college attendance, earnings or fertility behaviours (Chetty et al. (2014)). Yet, very little is known about what makes a teacher effective in raising student achievement. Since the estimation of teacher value-added is demanding in terms of data and generally requires the use of administrative datasets, most of the works trying to identify the determinants of teacher effectiveness has focused on teacher demographics and other observed characteristics, such as certification or tenure. Nevertheless, the literature fails to establish consistent and powerful relationships between teacher productivity measures and teacher observed characteristics (Aaronson et al. (2007)), with the notable exception of teacher experience, which is systematically related to higher levels of productivity[1].

This paper investigates the role of a largely unexplored and yet intuitive input of teacher productivity, namely the teaching practices she implements in the classroom. Exploiting US 8th grade students' data from the TIMSS 2011 assessment, I show that practices emphasizing student active participation in the lesson positively and strongly relate to math teachers' instructional productivity.

The TIMSS assessment encompasses 4 basic math topics (*Number*, *Algebra*, *Geometry* and *Data & Chance*) and each topic is divided into 3 to 6 subtopics (19 subtopics in total). For each teacher, the dataset provides information on the amount of instructional time devoted to each topic the year before the assessment as well as information on which subtopics were taught during this pre-assessment period. This wealth of data makes it possible to develop a strategy for identifying teachers' hourly productivity by focusing first on the performance of students on subtopics that were *not* taught the year before assessment and, second, on their performance on subtopics that were taught over this period.

Accordingly, when I first focus on subtopics that were *not* taught during the pre-assessment period, I find *no* relationship between the amount of instructional time devoted by teachers to the corresponding topics and the performance of students. This result is consistent with the assumption that the amount of instructional time devoted by teachers to a given topic is not related to students' initial level of ability in this specific topic.

Building on this assumption, I then provide estimates of teachers' hourly productivity by focusing on the subtopics that have been taught during the pre-assessment period and by looking at the relationship between students' performance in these subtopics and the amount of instructional time devoted by teachers to the corresponding topics. This analysis reveals that students' test scores in these subtopics strongly relate to the amount of instructional time devoted by teachers to the corresponding topic. Specifically, I

[1]See Harris & Sass (2011) for a summary of recent findings on that topic

find that a one hour increase in weekly instructional time in a given topic is associated with an average increase of about 4.4% of a SD in students' test scores on the corresponding subtopics.

In a last step, I investigate the extent to which estimated teachers' productivity levels relate to the teaching practices implemented in the classroom. Specifically, I explore the relationship between my measures of teachers' productivity and their use of practices emphasizing student active participation in the lessons, as opposed to teacher-centered practices and to practices based on student memorization and basic problem solving. I explore these issues with the aid of a "Modern Practices" index (MPI) constructed from the TIMSS survey.

Generally speaking, I find large productivity differentials across US math teachers according to the teaching practice they implement in the classroom. The effect of one additional weekly hour of instructional time on students' scores varies from 2.7% of a SD for teachers under the median of the MPI to 5.9% of a SD for teachers above the median of the MPI. Put differently, using the continuous specification of the MPI, I find that a one SD increase in this index relates to a 8% SD increase in test scores, which is roughly equivalent to half the effect of a SD increase in teacher value-added estimates from previous studies (Hanushek & Rivkin (2010)). An investigation of the potential mechanisms at play suggests that the positive effects associated to modern practices are partially mediated by an improvement in student self-confidence and motivation to learn mathematics.

This paper contributes to the small literature that explores the role of teaching practices in shaping teachers' effectiveness. Some recent papers provide evidence that pedagogical skills and the quality of student-teacher interactions strongly relate to teacher productivity (Kane et al. (2011), Blazar (2015) and Araujo et al. (2016)). In parallel, Machin & McNally (2008) and Lavy (2009) argue that the positive effects on student achievement generated by the *Literacy Hour* in the UK and a teacher payment scheme in Israel, respectively, were primarily mediated by changes in teaching methods. Altogether, these findings suggest that *teaching* - and not only *teachers* - may be a key determinant of instructional productivity, but they do not provide precise information on the teaching practices that are likely to improve teaching quality[2]. On the other hand, some recent papers has directly related *between subjects* or *between classes* variations in student test scores to variations in teaching practices across teachers, but they provide mixed results and do not give an insight into the magnitude of the relationship between teacher productivity and teaching practices[3]. The aim of this paper is to fill the gap between these two literatures by studying the relationship between the teaching practices implemented in the classroom by US math teachers and their instructional hourly productivity.

[2]This notwithstanding, it is important to note that all the measures of teaching quality used by Kane et al. (2011), Blazar (2015) and Araujo et al. (2016) emphasize the importance of student-teacher interactions

[3]These recent works include Aslam & Kingdon (2011), Schwerdt & Wuppermann (2011), Van Klaveren (2011), Bietenbeck (2014), Lavy (2015b) and Hidalgo-Cabrillana & Lopez-Mayan (2018)

This paper also contributes to the literature which aims at evaluating the causal effect of instructional time on student math test scores. This effect is an economically meaningful one, as student math skills have recently proven to be important predictors of both aggregate economic growth (Hanushek & Woessmann (2008) ; Hanushek & Woessmann (2011)) and individual's future earnings (Rose & Betts (2004) ; Joensen & Nielsen (2009) ; Goodman (2017)). Yet, there is only scarce evidence on this topic. Several recent papers find a positive impact on student test scores, but most of them rely on small variations or exploit programs that are targeted at specific students and generally accompanied with other changes in school's input[4]. Two notable exceptions are Lavy (2015a) and Rivkin & Schiman (2015), who both exploit *within student between subjects* variations in instructional time across countries, and rely on the assumption that these variations are independent from student subject specific-skills. Building on their work, this paper intends to improve the identification of the causal effect of instructional time through the exploitation of variations across topics of a single subject. This strategy arguably both requires less restrictive identification assumptions and allows for the exploitation of large variations.

The remainder of the paper is organized as follows. The next section describes the data and the construction of the teaching practices index. The second section presents the empirical strategy and provides some evidence on the validity of the identification assumptions. The third section presents the estimations of instructional productivity and its relationship with the teaching practices index. The final section concludes with a discussion of the implications of the main results.

1 The data

1.1 The TIMSS 2011 assessment

This paper exploits US data from the TIMSS 2011 assessment, which evaluates the math and science knowledge of eighth-grade students. The national sample is drawn from a two stage sampling procedure, whose objective is to ensure the national representativeness of US schools and students[5]. Every student in a selected class is assessed in math and science, and scores are assigned by independent external evaluators. This paper focuses on students' math test scores, which are important predictors of future earnings. The TIMSS math assessment encompasses 4 basic topics (*Number*, *Algebra*, *Geometry* and *Data and Chance*),

[4]Recent papers on this topic exploit variations in the number of school days over the year due to bad weather conditions or legal differences in the school start date (Sims (2008), Marcotte (2007) and Marcotte & Hemelt (2008)), remediation programs (Taylor (2014), Cortes et al. (2015)), or policy changes that increased resources allocated to schools, which result in an increased amount of instruction time (Bellei (2009), Lavy (2012) and Fryer (2014)). Two recent papers exploit a recent reform that took place in Germany and which implied a modest increase (+5%) in instructional time (Andrietti (2015) and Huebener et al. (2017))

[5]First, schools are randomly selected among the national sample of schools. In a second step, one class is selected in each selected school.

each of which is divided into 3 to 6 subtopics (19 subtopics in total). Finally, it is possible to compute a specific test score for each of these subtopics.

Besides the student assessment, every math teacher who teaches a selected class is asked to answer a questionnaire, which provides information on teacher demographics and teaching practices. I restrict the sample to students whose math teacher answered the teacher questionnaire, which amounts to dropping 30% of observations. The final sample is made up with 7258 students, allocated over 387 classes in 359 schools, and taught by 376 different teachers. The available evidence suggests that students in the final sample performed slightly better over the year than students whose math teacher didn't answer the questionnaire, though there doesn't seem to be large differences in terms of school and student characteristics according to teacher non response to the questionnaire[6].

1.2 Instructional time

The math teacher questionnaire includes detailed information about the total amount of instructional time that math teachers devote every week to each of the four basic topics in their class[7]. Importantly, students are taught these four topics in the same class, by the same teacher. As can be seen in table B3, students are given 4.4 hours of instructional time per week in math on average. Half of this time is spent on *Algebra*, and the rest is distributed in a more balanced fashion over the three remaining topics, though a smaller amount of time is devoted to *Data and Chance* on average ($\simeq$ 0.45 hour/week). In addition, there are substantial variations across teachers, both in the total amount of math instructional time per week and in the allocation of this time over the four topics. As argued in section 2, these observed variations in the share of instructional time devoted to the four topics might be mainly driven by the absence of a unique national curriculum in the US. Indeed, according to the TIMSS 2011 US National Research Coordinator, "the United States does not have a federally mandated national curriculum. State education agencies publish state mathematics standards and local school districts publish curriculum based on the standards"[8]. Such a variety of curricula introduces a lot of exogenous variations across schools and teachers in the allocation of instructional time across topics.

[6] As can be seen in tables B1 and B2 in the appendix, students in the final sample performed slightly better at the TIMSS assessment and are slightly older than those dropped from the initial sample due to teacher non response. No other difference appears to be significant between the two groups, regarding student and school characteristics.

[7] It is worth noting that the empirical strategy developed in this paper accounts for potential variations in the length of school year across schools that could introduce some measurement error in this measure of instructional time, as it is based on within student (and thus, within school) variations.

[8] Source: TIMSS Curriculum Questionnaire for Grade 8 (http://timssandpirls.bc.edu/timss2011/international-contextual-q.html)

1.3 Teaching practices

The measures of teaching practices used in this paper are drawn from question 19 in the math teacher questionnaire. For each of the 11 teaching practices listed in the questionnaire (cf. table 1), teachers are asked the following question: "In teaching math to the students in this class, how often do you usually ask them to do the following?". There are four possible answers to this question: "Every or almost every lesson", "About half the lessons", "Some lessons" or "Never". Table 1 exhibits the distribution of teachers' answers to this question for the different practices.

Table 1: **Definition and distribution of Teaching Practices**

Teaching practice ("I ask students to...")	Never	Some lessons	Half lessons	Every lesson
(a) Listen to me explain how to solve problems	1 (%)	16 (%)	16 (%)	67 (%)
(b) Memorize rules, procedures, facts	4	41	32	23
(c) Work pbs (individually or with peers) with my guidance	0	7	18	75
(d) Work pbs in whole class with direct guidance from me	1	12	20	67
(e) Work pbs (individually or with peers) while I am occupied	26	37	10	27
(f) Apply facts, concepts and procedures to solve routine pbs	0	14	24	62
(g) Explain their answers	0	12	27	61
(h) Relate what they learn to their daily lives	3	34	38	25
(i) Decide on their own procedure for solving complex pbs	3	36	35	26
(j) Work pbs for which there's no obvious method of solution	13	52	25	10
(k) Take a written test or quiz	0	58	25	17

Building on these questions, it is possible to construct for each practice and each teacher a measure of practice intensity, where intensity is set to 0 when the answer is "Never", to 1 when the answer is "Every or almost every lesson", 0.5 when the answer is "About half the lessons" and 0.1 when the answer is "Some lessons"[9]. To account for the fact that all teachers may not have the same definition of the different levels of intensity mentioned in the questionnaire (i.e., the same definition of "Every or almost every", for example) I also center these variable at teachers' means[10]. Overall, I obtain a set of variables describing the relative intensity of each practice for each teacher.

[9]Alternatively, I assign the score 0.25 to the answer "Some lessons" and check the robustness of my results to this alternative score. Results are presented in the section dedicated to robustness.

[10]Investigating relationships among self-declared practices in the dataset, I find that all pairwise correlation coefficients between teaching practices are positive or null (cf. table B4), which tends to support the existence of an individual bias in the way teachers answered these questions in the TIMSS survey.

2 The evaluation of instructional productivity

2.1 Estimation strategy

Assessing the causal impact of instructional time on student achievement raises two identification issues. First, schools with more fundings can both attract better teachers and students and give the latter a higher amount of instructional time, which would introduce an upward bias in the estimation of instructional productivity. Second, students could be assigned a better teacher and more instructional time based on their previous math achievement. This would introduce an upward or a downward bias, depending on the direction of this within school sorting. To overcome these issues, I exploit *within student* variations in math instructional time, which occur across math topics that are taught by the same teacher, at the same school. Formally, I estimate the following model:

$$A_{it} = \alpha_i + c_t + \beta_1 IT_{it} + \epsilon_{it} \tag{1}$$

where A_{it} is the TIMSS score in math topic $t \in \{1; 4\}$ of student i, and IT_{it} is the quantity of instructional time devoted to topic t by student i's math teacher. The model also includes student fixed effects (α_i), which captures student innate ability and motivation to learn mathematics. Importantly, as all math topics are taught by the same teacher at the same school for a given student, student fixed effects also include teacher and school fixed effects. To complete the model, I add topic-specific constants (c_t). Standard errors are systematically clustered at the teacher level.

The only determinants of student achievement that this specification does not control for are student math topic-specific skills. As a consequence, under the assumption that the *within student between topics* variations in instructional time (IT_{it}) are not related to student topic-specific skills (ϵ_{it}), β_1 identifies the causal effect of a weekly hour of instruction time on student test scores and thus provides a valid estimate of teachers' average hourly productivity.

The US educational system is characterized by the absence of a unique mathematics curriculum for 8th grade students. Based on the "state standard" published by the state education agency, each school district defines its own curriculum. As there are more than 14,000 school districts in the US, this system induces a lot of variations in the allocation of math instructional time across topics that is arguably exogenous to teachers and students.The main threat to this assumption is the possibility that, within the curriculum constraint, teachers adopt strategic behaviours which would consists in marginally allocating a higher (or lower) share of their instruction time to the topic in which their students perform relatively better (or worse).

To test the existence of teacher strategic behaviours that would bias the estimates, I take advantage of a particular feature of the TIMSS assessment. For each math topic under consideration, students are evaluated in both subtopics that are taught over the year preceding the test and subtopics that are *not* taught over this period[11]. Consequently, the test provides us with measures of students' topic-specific skills that are unaffected by the amount of instructional time that is dedicated to study the related topics over the year. Indeed, the instructional time devoted to a given topic the year of the test should positively affect student test scores in the related subtopics that are taught over the year, but not in the related subtopics that are *not* taught. Any relationship between the amount of instructional time devoted to a given topic and students' test scores in the related subtopics that are *not* taught over the year would instead capture teachers' strategic allocation of instructional time across math topics. Building on this argument, for all the estimations of instructional productivity presented in this paper, I implement regression (1) on the subtopics taught over the year only, and I show that there is no effect on subtopics *not* taught over the year.

2.2 US math teachers' instructional productivity

As can be seen in the first column of table 2, when considering subtopics that are taught the year of the assessment, I find that one weekly hour of instructional time increases student math test scores by 4,4% of a standard deviation on average, which roughly amounts to a 3,3 points increase in the TIMSS test score[12]. Contrarily, the coefficient associated to Instruction Time is not significant when considering student test scores in the subtopics that are not taught the year of the assessment (cf. column (2)). As discussed in the previous section, this tends to support the main identification assumption. This effect is quite large, compared with the effect of other school's input. For example, doubling the total amount of math instructional time would increase student test scores by 19.3% of a standard deviation over the year, while a 10 students reduction in class size would raise student test scores by 10 to 30% of a standard deviation, as estimated from previous studies (Hanushek & Rivkin (2010)). In addition, this estimation is consistent with previous studies investigating the effect of instruction time on student test scores in comparable settings[13].

[11]As previously mentioned, students are evaluated in 3 to 6 subtopics per topic (19 subtopics in total). Subtopics that have not been taught the year preceding the test may have been taught over previous years or have never been taught to the students taking the test.

[12]The mean test score in math in the final sample is 507

[13]In particular, studies evaluating the effect of mathematics instructional time in the US provide estimates ranging from 2.5% to 5% of a standard deviation (Dobbie & Fryer Jr (2013), Taylor (2014) and Cortes et al. (2015)). Other studies including Bellei (2009), Lavy (2012), Lavy (2015a), Rivkin & Schiman (2015) and Andrietti (2015) find an effect ranging from 2.1% to 7% of a standard deviation.

Table 2: Math teachers' instructional productivity

	(1) Subtopics taught	(2) Subtopics not taught
Instruction Time	0.044***	-0.001
	(0.010)	(0.008)
Observations	18888	22263

Note: this table shows the effect of one weekly hour of instructional time on student math test scores, separately for subtopics taught the year of the test (column (1)) and subtopics *not* taught the year of the test (column (2)). All regressions include student and teacher fixed effects, as well as topic constants. Standards errors (in parentheses) are clustered at the teacher level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3 Instructional productivity and teaching practices

Building on the estimates of math teachers' instructional productivity computed from within student variations in math instructional time, the second step of the empirical strategy consists in investigating the relationship between teachers' productivity and the teaching practices they implement in the classroom.

In order to describe the teaching style of a teacher in fewer dimensions than the 11 practices included in the questionnaire, I perform a Principal Component Analysis (PCA) at the teacher level. Based on this PCA, I create the *Modern Practices Index* (henceforth MPI), which equals teacher individual average score on practices (g), (h), (i), and (j). This index measures the relative importance of practices involving strong student-teacher interactions (practices (g) and (h)) and complex thinking (practices (i) and (j)) in the teaching style of the teacher, as opposed to teacher lecture ((a)) and basic problem-solving ((b), (c), (d), (e) and (f)), which are generally considered as traditional practices[14]. I complement this index with the frequency of assessment (practice (k)), which poorly relates to the MPI. Finally, I estimate the following model:

$$A_{it} = \alpha_i + c_t + \beta_1 IT_{it} + \beta_2.IT_{it}.MPI_i + \beta_3.IT_{it}.Assess_i + \epsilon_{it} \quad (2)$$

where MPI_i is the Modern Practices Index and $Assess_{ij}$ the frequency of assessment of student i's math teacher. The parameter β_2 indicates how teacher instructional productivity varies with the MPI. All other variables included in equation (2) are similar to those described in the previous paragraph for equation (1).

This strategy accounts for the potential endogeneity in the allocation of students to schools and teachers, as well as for the potential adaptation of teachers' teaching practices to the math general ability of the students in their class. Nevertheless, it is possible that the coefficient associated to the MPI reflects the effect of an unobserved teacher characteristics which is both related to the use of modern practices

[14] A detailed description of the results obtained from the PCA, as well as the construction of the MPI are available in section A of the appendix

and to student achievement. To mitigate this concern, I first show that the MPI is little influenced by the school and classroom environment, and that it is unrelated to teacher demographics (cf. table B5). By contrast, it is strongly and positively correlated with variables that relate to teachers' motivation and behavioral skills, such as collaborative behaviour and self-confidence[15]. Consequently, I sequentially add teacher characteristics as interacted controls in the regression:

$$A_{it} = \alpha_i + c_t + \beta_1 IT_{it} + \beta_2 .IT_{it}.MPI_i + \beta_3 .IT_{it}.Assess_i + \beta_4 .IT_{it}.X_i + \epsilon_{it} \tag{3}$$

where X_i is a vector of all teacher characteristics included in table B5, including teacher demographics and teacher behavioral controls. It also includes class size and the teacher perceived level of disruption in the classroom, which are two important determinants of instructional quality (Lazear (2001)). Results of the estimation of equation (3) are presented in table 3 and are discussed in the next section.

3.1 Math teachers' instructional productivity and Modern Practices

The use of practices emphasizing student active participation in the lesson is systematically associated to higher levels of teachers' instructional productivity. As can be seen in table 3, the coefficient associated to the interaction term between instructional time and the MPI is positive and significant in all specifications. In addition, this coefficient is remarkably stable across specifications. In particular, the inclusion of teacher behavioural controls, which strongly correlate to the MPI, has a very little impact on the MPI's estimated coefficient. This tends to support the idea that the MPI captures the quality of teaching and not solely the effect of some confounding factors such as teacher motivation[16]. In addition to this, the frequency of assessment is positively correlated to teachers' instructional productivity, though the corresponding coefficient is no longer significant when teacher behavioural controls are included in the regression.

To give insights about the magnitude of the variability in teacher productivity associated to the MPI, I provide two distinct interpretations. First, I examine how instructional productivity varies when moving

[15] Due to the absence of *within teacher* variations in teaching practices in the dataset, it is difficult to completely rule out the possibility that the MPI includes the effect of some confounding factors. On the whole, though adding teacher controls in the regression alleviates such a concern, one should be cautious regarding a causal interpretation of the effect of modern practices.

[16] To check the consistency of my results, I further check that the use of modern practices is unrelated to the allocation of instructional time across math topics. To do so, I regress the MPI of a given teacher on the percentages of instructional time she devotes to the different topics, controlling for the math average score of the students she teaches. Results are reported in table B6 in the appendix. As can be seen, none of the coefficients associated with the shares of instructional time devoted to the different topics is significant. In addition, I also compute the pairwise correlation coefficients between the MPI and the percentages of instructional time devoted to the topics. The only significant relationship that appears at the 10% level is a positive one between the share of instructional time devoted to *Geometry* and the MPI (cf. table B7). On the whole, there doesn't seem to be a strong relationship between the MPI and the allocation of instructional time across topics. Nevertheless, I check the robustness of my results to the exclusion of *Geometry* test scores (cf. table B16 in the appendix).

Table 3: Teaching Practices and Teachers' Instructional Productivity

	Score (1)	Score (2)	Score (3)	Score (4)	Score (5)
Panel A: subtopics taught (N=18888)					
Instructional Time (IT)	0.050***	0.061***	0.040	0.004	0.033
	(0.010)	(0.009)	(0.051)	(0.053)	(0.062)
IT*Modern Practices Index	0.096***	0.112***	0.108***	0.107***	0.099***
	(0.035)	(0.032)	(0.032)	(0.031)	(0.036)
IT*Assessment		0.050**	0.049**	0.042*	0.032
		(0.021)	(0.021)	(0.021)	(0.021)
Panel B: subtopics not taught (N=22263)					
Instructional Time (IT)	-0.001	-0.001	0.014	0.004	0.028
	(0.009)	(0.009)	(0.041)	(0.044)	(0.053)
IT*Modern Practices Index	0.005	0.005	0.012	0.017	0.006
	(0.024)	(0.024)	(0.024)	(0.025)	(0.025)
IT*Assessment		-0.001	-0.000	-0.001	-0.013
		(0.018)	(0.019)	(0.019)	(0.020)
IT*Teacher demographics	.	.	√	√	√
IT*Class size	.	.	.	√	√
IT*Teacher behaviour	.	.	.	.	√

Note: This table shows the heterogeneity in the effect of math instructional time on student math performance according to the teaching practices implemented in the classroom by the math teacher, separately on subtopics taught the year of the test (Panel A) and subtopics not taught the year of the test (Panel B). All regressions include student and teacher fixed effects, as well as topic constants and the proportion of subtopics taught the year of the test. Teacher demographic controls included in column (3) - (5) are teacher experience, gender and level of education and a dummy indicating if the teacher' major studied area was "Education-Mathematics". Controls included in column (5) include measures of teachers' collaboration with colleagues, self-confidence in teaching math and perceived level of disruption in the class drawn from the TIMSS teacher questionnaire and provided in the dataset, as well as a dummy indicating that the teacher participated in a professional development over the last two years. All controls are interacted with Instructional Time. Standards errors (in parentheses) are clustered at the teacher level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

along the MPI distribution. Assuming linearity in the effect of the MPI[17], I compute teacher instructional productivity at different points of the MPI distribution. Moving from the teacher at the 25th to the teacher at the 75th percentile of this distribution is equivalent to a 86% increase in instructional productivity (cf. table B8), which is substantial. Put differently, one hour of math instruction time spent with the latter teacher is about twice as productive as one hour spent with the former one. Second, I compute the effect of a standard deviation increase in the MPI on student test scores. A one standard deviation increase in the MPI increases student test scores by 0.018 of a standard deviation for each weekly hour of instruction time. Computing the effect for the whole year, a one standard deviation increase in the MPI increases student test scores by 0.08 of a standard deviation[18]. This is almost equivalent to doubling the total amount of instructional time, holding instructional productivity at its average level.

In a recent review, Hanushek & Rivkin (2010) show that the *teacher quality* literature provides estimates of the variability in teacher value-added that are highly consistent across studies. For mathematics teachers, a one standard deviation increase in teacher value-added is associated to a 0.15 standard deviation increase in student test scores over the year, on average. Similarly, moving from the teacher at the 25th percentile of the value-added distribution to the teacher at the 75th percentile during one single year is equivalent to a 0.2 standard deviation increase in math test scores. Using both interpretations, the MPI effect roughly equals half of the total teacher fixed effect (i.e. a one SD increase in the MPI equals half the effect of a SD increase in teacher value-added).

Furthermore, the magnitude of this effect is comparable to the results obtained by Kane et al. (2011) and Araujo et al. (2016), who use two distinct measures of pedagogical skills in order to assess the impact of teachers on US 3-8th grade and Ecuador 2-5th grade student achievement, respectively. In the first case, a one standard deviation in the TES score, which measures the quality of student-teacher relations and teacher global instructional skills through the assessment of an external evaluator, increases student test scores by 0.05 standard deviation. In the second case, a one standard deviation increase in the CLASS score, which measures the quality of teacher behaviours in terms of emotional support, classroom organization and instructional support through video observations, increases student test scores by 0.06-0.09 standard deviation over the year, depending on the specification. Importantly, both these measures put a high weight on the quality of student-teacher interactions, which is also the case for the MPI. This tends to confirm the idea that these interactions are crucial in shaping teachers' instructional productivity.

[17] I investigate the extent to which the relationship between instructional productivity and the MPI is linear in the robustness checks section

[18] The effect of a SD increase in the MPI on hourly productivity is computed as follows: $\hat{\sigma}_{MPI} * \hat{\beta}_2 = 0.189 * 0.096 = 0.018$. Over the year, the effect is 4.4*0.018=0.079

3.2 Robustness checks and heterogeneity analysis

The main results outlined in this paper are robust to several alternative specifications regarding the definition of teaching practice variables.

First, I provide evidence that the way I assign scores to the teaching practice variables in the main specification is not driving the results. As can be seen in tables B9 and B10, respectively, considering a binary definition of teaching practice variables[19] or assigning the score 0.25 to the answer "Sometimes" leads to the same conclusion.

Second, I show that the main results are not driven by the correction applied to teaching practice variables, which objective is to take into account individual biases in teachers' answers. To explore this issue, I construct one Modern Practices Index and one Traditional Practices Index based on the non centered values of teaching practices, and I include both indexes in the regression[20]. As can be seen in table B12, this specification gives similar results. Indeed, the "non centered" MPI is strongly and positively associated to instructional productivity while the coefficient associated to the "non centered" Traditional Practices Index is negative, though it's not significant at conventional confidence levels.

Third, the conclusions drawn from the main specification are robust to considering more dimensions of teaching practices than those captured by the MPI and the frequency of assessment. Indeed, including the teacher total score on all practices to take into account the diversity of practices leaves the coefficient associated to the MPI roughly unchanged (cf. table B13). Furthermore, including the two traditional practices indexes described in section 1 instead of the single MPI also leads to the same conclusion, as the coefficients associated to both indexes are strongly significant and negative (cf. table B14)[21].

Fourth, the magnitude of the effect associated to the use of modern practices is unchanged when comparing the productivity of teachers who rank in the bottom half of the MPI distribution vs the top half, instead of using a continuous definition of the MPI. Indeed, teachers in the top half of the MPI distribution have an average productivity of 0.059 σ-test score per weekly hour, which is twice as large as the productivity of teachers who belong to the top bottom of this distribution (cf. table B15)[22].

In addition to these robustness checks regarding the specification of teaching practice variables, I also check that the results obtained from the main specification are not driven by the inclusion of *Geometry*

[19] In the binary model, the score 1 is assigned to the answer "At every lesson" and 0 to the three other answers. The estimated coefficients from this regression are smaller and less significant than those obtained from the main regressions, as considering a binary definition of teaching practice variables amounts to lose a lot of information.

[20] In the main specification, the centered Modern Practices Index is strongly and negatively related to the Traditional Practices index (cf. table B11). By contrast, when computed over non centered values of teaching practice variables, these two indexes exhibit a small and positive correlation coefficient of 0.12. Consequently, both indexes are included in the regression.

[21] This specification better accounts for the second dimension of teaching practices highlighted in the principal component analysis.

[22] Unfortunately, the sample size is too small to precisely estimate teachers' instructional productivity at different points of the MPI distribution when the number of categories is higher than 2.

test scores. As can be seen in table B16, these results are robust to the exclusion of *Geometry* test scores from the regression, as it doesn't affect the coefficients associated to the MPI.

Finally, I investigate whether the MPI effect differs by student gender. Implementing equation (3) separately on girls and boys, I find no significant differences in the coefficient associated to the MPI (cf. table B17).

3.3 Potential mechanisms: Modern Practices and student non cognitive outcomes

This section investigates the extent to which the positive effect associated to the use of modern practices is mediated by an improvement of student non cognitive outcomes. Three measures of non cognitive outcomes are available in the dataset: student self-confidence in learning mathematics and student intrinsic and extrinsic motivation to learn mathematics[23]. As these outcomes are measured at the end of the year, they are plausibly affected by the teachers observed in the dataset and the teaching practices they have implemented in the classroom over the year.

As there is no within student variations in non cognitive outcomes in the dataset, I estimate the relationship between the MPI and student non cognitive outcomes through the following model:

$$NCO_{ij} = \alpha + \beta_1 MPI_{ij} + \beta_2 A_{i0} + \beta_3 X_{ij} + \epsilon_{ij} \tag{4}$$

where NCO_{ij} is the non cognitive outcome score of student i, taught by teacher j. MPI_{ij} is the Modern Practices Index of teacher j and A_{i0} is student i's math mean score, computed over subtopics *not* taught over the year, that are presumably unaffected by the teaching practices implemented by the observed teacher. This *proxy* for student math ability intends to control for the fact that initially better students, who also have better non cognitive outcomes, could be assigned teachers who rank higher on the MPI. Finally, X_{ij} is a vector of controls including student gender, age, socio-economic background and language spoken at home, as well as the amount of math instructional time per week, school size, indexes of school immediate area's economic affluence and urban density and all teacher characteristics included in equation (3).

As can be seen in table B18, the use of modern practices is positively associated to the three non cognitive outcomes under consideration, though the relationship is not significant for student self-confidence at conventional levels. This result is consistent with the notion that the use of modern practices leads students to engage more actively in mathematics lesson. This attitude may, in turn, help them improve their math performance. Furthermore, this result is consistent with Algan et al. (2013), who find that

[23]These measures are drawn from questions 14 and 16 in the student questionnaire, and are directly provided in the database. A detailed description of the construction of these measures is available on the TIMSS website: https://timssandpirls.bc.edu/methods/t-context-q-scales.html

teaching practices which imply strong student-teacher interactions and interactions among students are associated with higher levels of self-confidence and positive attitudes toward learning mathematics.

4 Conclusion

The results outlined in this paper shed a new light on the determinants of teacher instructional productivity and the mechanisms lying behind the large heterogeneity observed across US teachers. Building on a new empirical strategy to estimate teachers hourly productivity, I show that the use of practices emphasizing student active participation in the lesson is systematically associated with higher levels of productivity. Specifically, I construct an index measuring the relative weight that math teachers put on these practices and I show that teachers above this index's median are twice as productive as teachers under the median. In terms of magnitude, I find that a one SD increase in this index is related to a 0.08 SD increase in student test scores over the year, which is equivalent to half the effect of a SD increase in teacher value-added estimates from previous studies. A further investigation of the potential mechanisms at play suggests that this effect is mediated by an increase in student self-confidence and motivation to learn mathematics.

These results confirm that teachers are a key determinant of student achievement and suggest a new way to improve teachers' productivity through the promotion of better teaching practices. An important area for future research is to determine the extent to which the positive relationship between teacher instructional productivity and practices based on student active participation truly reflects the causal effect of these practices. In particular, it is possible that only teachers endowed with a high level of pedagogical skills are able to efficiently implement these practices. In this case, forcing teachers (including those poorly endowed with pedagogical skills) to implement them could be counterproductive. In addition, it is important to take into account the adjustment costs incurred by a policy aiming at enhancing new practices, as teachers are not necessarily able to instantaneously absorb and retain new teaching methods. To my knowledge, the only paper dealing with these issues is the one by Haeck et al. (2014) who study the effect of a universal school reform implemented in the early 2000's in Quebec. Their findings are consistent with the existence of adjustment costs and therefore confirm that investigating the long term cost effectiveness of such policies in a dynamic and experimental setting is a key area for future research.

References

Aaronson, D., Barrow, L., & Sander, W. 2007. "Teachers and student achievement in the Chicago public high schools". *Journal of Labor Economics*, 25(1):95–135.

Algan, Y., Cahuc, P., & Shleifer, A. 2013. "Teaching Practices and Social Capital". *American Economic Journal: Applied Economics*, 5(3):189–210.

Andrietti, V. 2015. "The causal effects of increased learning intensity on student achievement: Evidence from a natural experiment". UC3M Working Paper Economic Series 15-06.

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. 2016. "Teacher Quality and Learning Outcomes in Kindergarten". *The Quarterly Journal of Economics*, 131(3):1415–1453.

Aslam, M. & Kingdon, G. 2011. "What can teachers do to raise pupil achievement?". *Economics of Education Review*, 30(3):559–574.

Bellei, C. 2009. "Does lengthening the school day increase students academic achievement? Results from a natural experiment in Chile". *Economics of Education Review*, 28(5):629–640.

Bietenbeck, J. 2014. "Teaching practices and cognitive skills". *Labour Economics*.

Blazar, D. 2015. "Effective teaching in elementary mathematics: Identifying classroom practices that support student achievement". *Economics of Education Review*, 48:16–29.

Bruner, J. S. 1961. "The act of discovery.". *Harvard educational review*.

Chetty, R., Friedman, J. N., & Rockoff, J. E. 2014. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood". *The American Economic Review*, 104(9):2633–2679.

Cortes, K. E., Goodman, J. S., & Nomi, T. 2015. "Intensive math instruction and educational attainment long-run impacts of double-dose algebra". *Journal of Human Resources*, 50(1):108–158.

Dobbie, W. & Fryer Jr, R. G. 2013. "Getting beneath the veil of effective schools: Evidence from New York City". *American Economic Journal: Applied Economics*, 5(4):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(3):1355–1407.

Goodman, J. 2017. "The Labor of Division: Returns to Compulsory High School Math Coursework". National Bureau of Economic Research.

Haeck, C., Lefebvre, P., & Merrigan, P. 2014. "The distributional impacts of a universal school reform on mathematical achievements: A natural experiment from Canada". *Economics of Education Review*, 41:137–160.

Hanushek, E. A. & Rivkin, S. G. 2006. "Teacher quality". In *Handbook of the Economics of Education*, volume 2, pages 1051–1078. Elsevier, 2006.

Hanushek, E. A. & Rivkin, S. G. 2010. "Generalizations about using value-added measures of teacher quality". *The American Economic Review*, 100(2):267–271.

Hanushek, E. A. & Woessmann, L. 2008. "The role of cognitive skills in economic development". *Journal of Economic Literature*, 46(3):607–668.

Hanushek, E. A. & Woessmann, L. 2011. "The Economics of International Differences in Educational Achievement". In *Handbook of the Economics of Education*, volume 3, pages 89–200. Elsevier, 2011.

Harris, D. N. & Sass, T. R. 2011. "Teacher training, teacher quality and student achievement". *Journal of Public Economics*, 95(7):798–812.

Hidalgo-Cabrillana, A. & Lopez-Mayan, C. 2018. "Teaching styles and achievement: Student and teacher perspectives". *Economics of Education Review*, 67:184–206.

Huebener, M., Kuger, S., & Marcus, J. 2017. "Increased instruction hours and the widening gap in student performance". *Labour Economics*.

Joensen, J. S. & Nielsen, H. S. 2009. "Is there a causal effect of high school math on labor market Outcomes?". *Journal of Human Resources*, 44(1):171–198.

Kane, T. J., Taylor, E. S., Tyler, J. H., & Wooten, A. L. 2011. "Identifying effective classroom practices using student achievement data". *Journal of Human Resources*, 46(3):587–613.

Lavy, V. 2009. "Performance pay and teachers' effort, productivity, and grading ethics". *The American Economic Review*, 99(5):1979–2021.

Lavy, V. 2012. "Expanding school resources and increasing time on task: effects of a policy experiment in Israel on student academic achievement and behavior". National Bureau of Economic Research.

Lavy, V. 2015a. "Do differences in schools' instruction time explain international achievement gaps? Evidence from developed and developing countries". *The Economic Journal*, 125(588):F397–F424.

Lavy, V. 2015b. "What makes an effective teacher? Quasi-experimental evidence". *CESifo Economic Studies: Oxford Journals*.

Lazear, E. P. 2001. “Educational Production”. *The Quarterly Journal of Economics*, 116(3):777–803.

Machin, S. & McNally, S. 2008. “The literacy hour”. *Journal of Public Economics*, 92(5):1441–1462.

Marcotte, D. E. 2007. “Schooling and test scores: A mother-natural experiment”. *Economics of Education Review*, 26(5):629–640.

Marcotte, D. E. & Hemelt, S. W. 2008. “Unscheduled school closings and student performance”. *Education*, 3(3):316–338.

Mayer, R. E. 2004. “Should there be a three-strikes rule against pure discovery learning?”. *American psychologist*, 59(1):14.

National Council of Teachers of Mathematics. 1991. *Professional standards for teaching mathematics.* Commission on Teaching Standards for School Mathematics.

Piaget, J. 1970. *Science of education and the psychology of the child.* Trans. D. Coltman. Orion.

Rivkin, S. G. & Schiman, J. C. 2015. “Instruction time, classroom quality, and academic achievement”. *The Economic Journal*, 125(588):F425–F448.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. 2005. “Teachers, schools, and academic achievement”. *Econometrica*, 73(2):417–458.

Rockoff, J. E. 2004. “The impact of individual teachers on student achievement: Evidence from panel data”. *American Economic Review*, pages 247–252.

Rose, H. & Betts, J. R. 2004. “The effect of high school courses on earnings”. *Review of Economics and Statistics*, 86(2):497–513.

Schwerdt, G. & Wuppermann, A. C. 2011. “Is traditional teaching really all that bad? A within-student between-subject approach”. *Economics of Education Review*, 30(2):365–379.

Sims, D. P. 2008. “Strategic responses to school accountability measures: It’s all in the timing”. *Economics of Education Review*, 27(1):58–68.

Taylor, E. 2014. “Spending more of the school day in math class: Evidence from a regression discontinuity in middle school”. *Journal of Public Economics*, 117:162–181.

Van Klaveren, C. 2011. “Lecturing style teaching and student performance”. *Economics of Education Review*, 30(4):729–739.

Vygotsky, L. S. 2012. *Thought and language.* MIT press.

Appendix A Principal Component Analysis and construction of the Modern Practices Index

This appendix describes the construction of the *Modern Practices Index*, which is the main measure of teaching practices used in this paper. I first perform a Principal Component Analysis (PCA) at the teacher level, including the 11 teaching practice variables described in section 1.3. Figure A1 plots the different practices on the two first axis of the PCA, which summarizes 37% of the between teacher total variation in these 11 variables. The first axis clearly opposes *student-centered* practices, which are based on student active participation, to *teacher-centered* practices and practices based on memorization and routine problems solving. These two sets of practices roughly correspond to what has been called *Modern practices* and *Traditional practices* in the economics of education literature, and this classification is consistent with the main psychological theories of learning. In particular, these theories oppose the *transmissive* approach, where the teacher delivers knowledge to a passive learner, and the *constructivist* or *socio-constructivist* approach[24], which has been promoted in the US by the National Council of Teachers of Mathematics (1991) over the last two decades and for which "learning is an active process in which learners are active sense makers who seek to build coherent and organized knowledge" (Mayer (2004)).

To sum up the opposition between the two sets of practices, I create the *Modern Practices Index* (MPI), which is equal to the individual teacher's average score over practices (g), (h), (i) and (j). The MPI goes from -0.5 and 0.5, with a mean of -0.07 (cf. table A1), and is roughly normally distributed (cf. figure A2). In order to take into account the second axis of the PCA, the frequency of assessment (practice (k)) is included separately in the regressions. I additionally create two *Traditional* indexes corresponding to the two subsets of traditional practices, in order to check the robustness of my results to considering more dimensions of teaching. These two indexes equal the teacher's average score over practices (a), (c) and (d), and practices (b), (e) and (f), respectively.

Table A1: Distribution of the Modern Practices Index

Variable	Mean	SD	Min	p25	p50	p75	Max
Modern Practices Index	-0.07	0.19	-0.51	-0.21	-0.08	0.06	0.49

Note: This table shows the mean and standard deviation of the Modern Practices Index (MPI). It also shows the minimum, the maximum, and the quartiles (p25, p50 and p75) of the MPI.

[24] See Piaget (1970), Bruner (1961) and Vygotsky (2012)

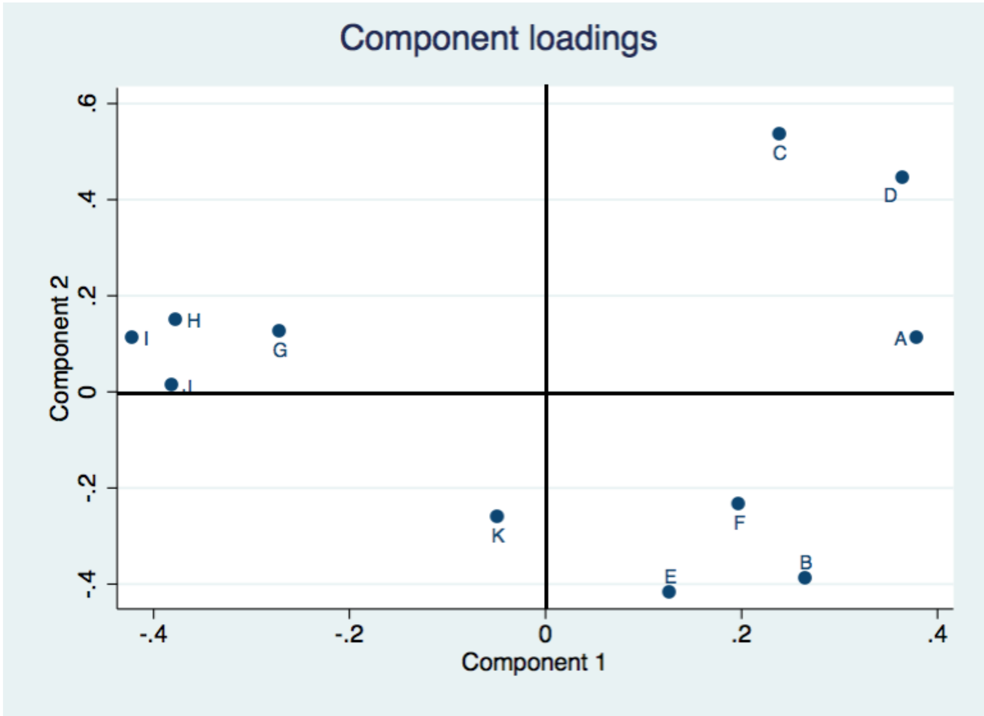

Figure A1: Principal Component Analysis - Teaching Practices

Note: Figure A1 plots the component loadings of the 11 teaching practices listed in table 1 on the two first axis of the principal component analysis, which is performed at the teacher level. Teaching practices are denoted with a letter, which refers to table 1.

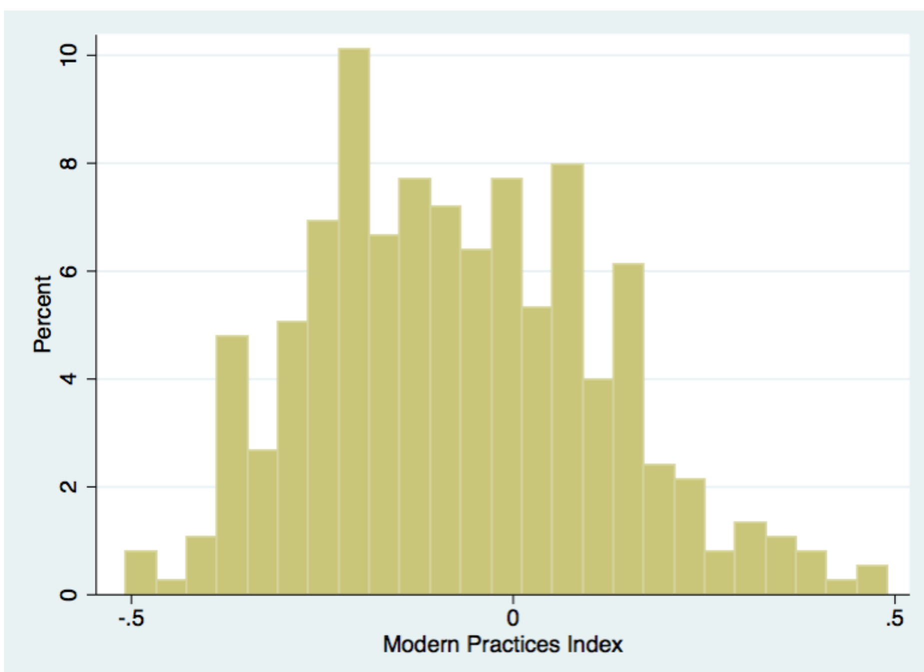

Figure A2: Distribution of the Modern Practices Index

Appendix B Additional Tables and Graphs

Table B1: Teacher non response and student characteristics

Variable	Final sample (1)	Dropped students (2)	Mean Difference (1) - (2)
Female	0.508	0.493	0.0128 (0.89)
Age	14.26	14.22	0.0361* (1.82)
Foreign language spoken at home	1.374	1.387	-0.0122 (-0.35)
Educational aspirations	5.303	5.263	0.0434 (1.14)
Nb of books at home	2.882	2.884	0.00379 (0.06)
Parents' education	2.033	2.016	0.0171 (0.27)
Math test score	507	496	11* (1.95)
N	372	163	

Note: This table shows the mean characteristics of students whose math teachers answered the teacher questionnaire (column (1)) and students whose math teacher didn't answer the questionnaire (column (2)), in terms of student age, gender, language spoken at home, educational aspirations, parental education and math performance at the TIMSS test, computed at the teacher level. Eventually, column (3) shows the difference between these two groups of students and provides t-test of the significance of the average difference in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B2: Teacher non response and school characteristics

Variable	Final sample (1)	Dropped schools (2)	Mean Difference (1) - (2)
School size	727	740	-13
			(-0.37)
School remoteness	3.51	3.26	0.254
			(1.53)
Average income level of area	2.333	2.272	0.06
			(0.89)
Total number of computers	116.28	118.46	-2.175
			(-0.22)
Shortage of math teacher	1.559	1.512	0.046
			(0.48)
Math resource shortages	11.02	10.96	0.051
			(0.20)
N	329	125	456

Note: This table shows the mean characteristics of schools in which the math teachers answered the teacher questionnaire (column (1)) and schools in which the math teacher didn't answer the questionnaire (column (2)), in terms of school size, remoteness, average income level of area, number of computers and math teachers' and resources' shortages, computed at the school level. Eventually, column (3) shows the difference between these two groups of schools and provides t-test of the significance of the average difference in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B3: Distribution of math instructional time across math topics

	Mean	SD	p25	p50	p75
Math instructional time in hours/week:					
Number	0.85	0.73	0.35	0.72	1.15
Algebra	2.37	1.44	1.26	2.22	3.33
Geometry	0.75	0.79	0.21	0.58	1.00
Data & Chance	0.45	0.38	0.19	0.39	0.66
Total	4.42	1.63	3.75	4.17	5.00

Note: This table shows the mean and standard deviation of math instructional time per topics, expressed in hours per week and computed at the teacher level. $p25$, $p50$ and $p75$ respectively represent the 25th, the 50th and the 75th percentile of the instructional time variable distribution.

Table B4: Pairwise correlation coefficients among teaching practice variables

Teaching Practice	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)
(a)	1.000										
(b)	0.364**	1.000									
(c)	0.334**	0.242**	1.000								
(d)	0.486**	0.324**	0.495**	1.000							
(e)	0.200**	0.274**	0.215**	0.202**	1.000						
(f)	0.346**	0.377**	0.358**	0.316**	0.327**	1.000					
(g)	-0.003	0.046	0.107**	0.072	0.065	0.238**	1.000				
(h)	-0.078	0.001	0.048	-0.013	-0.014	0.090*	0.277**	1.000			
(i)	-0.055	0.032	0.118**	-0.002	0.127**	0.144**	0.284**	0.406**	1.000		
(j)	0.001	0.083	0.098*	-0.003	0.129**	0.127**	0.270**	0.295**	0.509**	1.000	
(k)	0.092*	0.164**	0.034	0.041	0.024	0.142**	0.064	0.106**	0.019	0.099*	1.000

Note: This table shows all the pairwise correlation coefficients among teaching practice variables. Teaching practices are denoted with a letter, which refers to table 1. * $p < 0.10$, ** $p < 0.05$.

Table B5: Modern Practices Index and Teacher, School and Student characteristics

	Correlation coefficient
Teacher characteristics (N=372)	
Experience	-0.06
Female	0.03
Education level	-0.01
Major area of study = mathematics	0.05
Major area of study = education - mathematics	0.11**
Professional development in math content	0.18***
Professional development in math pedagogy	0.13**
Professional development in math curriculum	0.11**
Confidence in teaching math	0.38***
Collaboration with colleagues	0.13**
School characteristics (N=355)	
School size	-0.01
School remoteness	-0.05
Average income level of area	0.03
Total number of computers	-0.05
Math resource shortages	-0.03
Student and class characteristics (N=372)	
Female	0.09*
Age	0.05
Foreign language spoken at home	0.06
Educational aspirations	0.04
Nb of books at home	-0.06
Parents' education level	-0.08
Class size	-0.03
Classroom disruption (perceived by the teacher)	-0.07

Note: This table shows pairwise correlation coefficients between the Modern Practices Index (MPI) and teacher, school and student characteristics. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$.

Table B6: Modern Practices Index and the allocation of Instructional Time across topics - regression

	(1) Modern Practices Index
Number	-0.0019
	(0.0012)
Algebra	-0.0014
	(0.0010)
Geometry	-0.0003
	(0.0012)
Data & chance	0.0008
	(0.0018)
Observations	372

Note: This table shows the estimated coefficients from the regression of the Modern Practices Index (MPI) on the percentages of instructional time devoted to each of the four math topics, controlling for the class mean score in math, computed over subtopics not taught the year of the TIMSS assessment. The regression is implemented at the teacher level. Standard errors are in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B7: Modern Practices Index and the allocation of Instructional Time across topics - pairwise correlation coefficients

	Correlation coefficient
Number	-0.06
Algebra	-0.07
Geometry	0.09*
Data & chance	0.07
Observations	372

Note: This table shows the correlation coefficients between the Modern Practices Index (MPI) and the percentage of math instructional time dedicated to each of the four topics. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B8: Instructional Productivity along the MPI distribution

Position of the teacher in the *MPI* distribution (1)	MPI_{pth} value (2)	Teacher Instructional Productivity (in σ-test score) (3)	Change in productivity relative to the median teacher (4)
10th percentile	-0.30	0.021	- 50%
25th percentile	-0.21	0.030	- 28%
50th percentile	-0.08	0.042	0%
75th percentile	0.06	0.056	+ 33%
90th percentile	0.16	0.065	+ 55%

Note: This table shows the effect of one weekly hour of math instructional time on student performance in math, estimated at different points of the *Modern Practices Index* (MPI) distribution, assuming a linear relationship bewteen the MPI and teachers' instructional productivity. Point estimates shown in column (3) are computed as follows: $Productivity_{pth} = \hat{\beta}_1 + \hat{\beta}_2 MPI_{pth}$, with MPI_{pth} the value of MPI in column (2) and $\hat{\beta}_1$ and $\hat{\beta}_2$ the coefficients associated to Instructional Time and to the interaction term between Instructional Time and MPI, respectively, estimated from the main regression. The first line of the Table might be interpreted as follows: one weekly hour of instructional time given by the teacher at the 10th percentile of the MPI distribution increases student test scores by 2.1% of a standard deviation, which is 50% less productive than one hour taught by the teacher at the median of the MPI distribution.

Table B9: Robustness check - Different score for Teaching Practice variables

	Score (1)	Score (2)	Score (3)	Score (4)	Score (5)
Panel A: subtopics taught (N=18888)					
Instructional Time	0.050***	0.061***	0.040	0.004	0.033
	(0.010)	(0.009)	(0.051)	(0.053)	(0.062)
IT*Modern Practices 2	0.096***	0.112***	0.108***	0.107***	0.099***
	(0.035)	(0.032)	(0.032)	(0.031)	(0.036)
IT*Assessment		0.050**	0.049**	0.042*	0.032
		(0.021)	(0.021)	(0.021)	(0.021)
Panel B: subtopics not taught (N=22263)					
Instructional Time	-0.001	-0.001	0.014	0.004	0.028
	(0.009)	(0.009)	(0.041)	(0.044)	(0.053)
IT*Modern Practices 2	0.005	0.005	0.012	0.017	0.006
	(0.024)	(0.024)	(0.024)	(0.025)	(0.025)
IT*Assessment		-0.001	-0.000	-0.001	-0.013
		(0.018)	(0.019)	(0.019)	(0.020)
IT*Teacher demographics	.	.	√	√	√
IT*Class size	.	.	.	√	√
IT*Teacher behaviour	.	.	.	.	√

Note: This table replicates table 3, using an alternative definition of the Modern Practices Index (MPI), which is based on a different way of scoring teachers' answers to the questions related to teaching practices. To compute this alternative MPI, I assign the score 0.25 (instead of 0.1) to the answer "sometimes" for all teaching practices variables. All regressions include student and teacher fixed effects, as well as topic constants and the proportion of subtopics taught the year of the test. Controls included in columns (3) - (5) are similar to those described in table 3. Standards errors (in parentheses) are clustered at the teacher level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B10: Robustness check - Binary Teaching Practice variables

	Score (1)	Score (2)	Score (3)	Score (4)	Score (5)
Panel A: subtopics taught (N=18888)					
Instructional Time (IT)	0.052***	0.059***	0.029	-0.008	0.032
	(0.011)	(0.011)	(0.054)	(0.056)	(0.065)
IT*Modern Practices (binary)	0.060*	0.063*	0.057*	0.056*	0.050
	(0.034)	(0.034)	(0.033)	(0.033)	(0.034)
IT*Assessment (binary)		0.027	0.028	0.022	0.015
		(0.018)	(0.017)	(0.017)	(0.018)
Panel B: subtopics not taught (N=22263)					
Instructional Time (IT)	-0.001	-0.001	0.015	0.006	0.029
	(0.009)	(0.010)	(0.041)	(0.044)	(0.054)
IT*Modern Practices (binary)	0.002	0.002	0.007	0.009	0.002
	(0.020)	(0.020)	(0.021)	(0.022)	(0.021)
IT*Assessment (binary)		-0.001	-0.000	-0.000	-0.004
		(0.014)	(0.015)	(0.015)	(0.015)
IT*Teacher demographics	.	.	√	√	√
IT*Class size	.	.	.	√	√
IT*Teacher behaviour	.	.	.	.	√

Note: This table replicates table 3, using an alternative definition of the Modern Practices Index (MPI), which is based on a categorical definition of teaching practice variables. To compute this alternative MPI, I assign the score 1 to the answer "Every or almost every lesson" and 0 to all other answers. All regressions include student and teacher fixed effects, as well as topic constants and the proportion of subtopics taught the year of the test. Controls included in columns (3) - (5) are similar to those described in table 3. Standards errors (in parentheses) are clustered at the teacher level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B11: Pairwise correlation coefficients among Teaching Practices Indexes

	MPI	TPI_1	TPI_2	TPI_3
Modern Practices Index (MPI)	1			
Traditional Practices Index 1 (TPI_1)	-0.912***	1		
Traditional Practices Index 2 (TPI_2)	-0.650***	0.712***	1	
Traditional Practices Index 3 (TPI_3)	-0.612***	0.672***	-0.042	1

Note: this table exhibits pairwise correlation coefficients between the Modern Practices Index and the Traditional Practices Indexes. TPI_1 includes all the 6 traditional practices, whereas TPI_2 only include practices (a), (c) and (d) and TPI_3 only include practices (b), (e) and (f), respectively. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$.

Table B12: Robustness check - Non centered Value of Teaching Practices

	Score (1)	Score (2)	Score (3)	Score (4)	Score (5)
Panel A: subtopics taught (N=18888)					
Instructional Time (IT)	0.039	0.027	-0.001	-0.044	-0.012
	(0.027)	(0.026)	(0.060)	(0.063)	(0.072)
IT*Traditional Practices Index'	-0.053*	-0.035	-0.031	-0.027	-0.028
	(0.030)	(0.030)	(0.031)	(0.031)	(0.031)
IT*Modern Practices Index'	0.077**	0.101***	0.104***	0.107***	0.100***
	(0.035)	(0.033)	(0.034)	(0.033)	(0.035)
IT*Assessment		0.050**	0.050**	0.044**	0.035*
		(0.022)	(0.021)	(0.021)	(0.021)
Panel B: subtopics not taught (N=22263)					
Instructional Time (IT)	0.001	0.001	0.022	0.011	0.046
	(0.023)	(0.023)	(0.042)	(0.047)	(0.059)
IT*Traditional Practices Index'	-0.004	-0.005	-0.017	-0.017	-0.021
	(0.024)	(0.024)	(0.026)	(0.026)	(0.026)
IT*Modern Practices Index'	0.002	0.002	0.001	0.005	-0.010
	(0.019)	(0.019)	(0.018)	(0.020)	(0.022)
IT*Assessment		-0.002	-0.004	-0.005	-0.020
		(0.019)	(0.020)	(0.020)	(0.022)
IT*Teacher demographics	.	.	√	√	√
IT*Class size	.	.	.	√	√
IT*Teacher behaviour	.	.	.	.	√

Note: This table replicates table 3, using two distinct teaching practices indexes, one Traditional and one Modern, computed over the non centered values of teaching practice variables. All regressions include student and teacher fixed effects, as well as topic constants and the proportion of subtopics taught the year of the test. Controls included in columns (3) - (5) are similar to those described in table 3. Standards errors (in parentheses) are clustered at the teacher level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B13: Robustness check - Including the diversity of Teaching Practices

	Score (1)	Score (2)	Score (3)	Score (4)	Score (5)
Panel A: subtopics taught (N=18888)					
Instructional Time (IT)	0.029	0.027	-0.001	-0.044	-0.012
	(0.027)	(0.026)	(0.060)	(0.063)	(0.072)
IT*Modern Practices Index	0.102***	0.124***	0.125***	0.125***	0.119***
	(0.037)	(0.034)	(0.034)	(0.033)	(0.037)
IT*Teaching Practices Diversity	0.004	0.006	0.007	0.007	0.007
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
IT*Assessment		0.056***	0.055***	0.049**	0.039*
		(0.021)	(0.021)	(0.021)	(0.021)
Panel B: subtopics not taught (N=22263)					
Instructional Time (IT)	0.001	0.001	0.022	0.011	0.046
	(0.022)	(0.023)	(0.042)	(0.047)	(0.059)
IT*Modern Practices Index	0.005	0.005	0.012	0.016	0.004
	(0.024)	(0.024)	(0.024)	(0.025)	(0.025)
IT*Teaching Practices Diversity	-0.000	-0.000	-0.001	-0.001	-0.003
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
IT*Assessment		-0.001	-0.001	-0.002	-0.016
		(0.018)	(0.019)	(0.019)	(0.021)
IT*Teacher demographics	.	.	√	√	√
IT*Class size	.	.	.	√	√
IT*Teacher behaviour	.	.	.	.	√

Note: This table replicates table 3, using an index of teaching practice diversity in addition to the main Modern Practices Index. The index of diversity equals the total score of the teacher on the 11 teaching practices. All regressions include student and teacher fixed effects, as well as topic constants and the proportion of subtopics taught the year of the test. Controls included in columns (3) - (5) are similar to those described in table 3. Standards errors (in parentheses) are clustered at the teacher level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B14: Robustness check - Two distinct Traditional Practices Indexes

	Score (1)	Score (2)	Score (3)	Score (4)	Score (5)
Panel A: subtopics taught (N=18888)					
Instructional Time (IT)	0.049***	0.052***	0.034	0.001	0.031
	(0.012)	(0.012)	(0.049)	(0.052)	(0.061)
IT*Traditional Practices Index 2 (TPI_2)	-0.067**	-0.062**	-0.060*	-0.063**	-0.052
	(0.032)	(0.031)	(0.031)	(0.030)	(0.033)
IT*Traditional Practices Index 3 (TPI_3)	-0.114***	-0.109***	-0.106***	-0.102***	-0.101**
	(0.038)	(0.039)	(0.040)	(0.039)	(0.041)
IT*Assessment		0.022	0.022	0.016	0.007
		(0.022)	(0.021)	(0.022)	(0.022)
Panel B: subtopics not taught (N=22263)					
Instructional Time (IT)	0.004	0.004	0.020	0.009	0.035
	(0.010)	(0.010)	(0.041)	(0.043)	(0.052)
IT*Traditional Practices Index 2 (TPI_2)	-0.023	-0.023	-0.023	-0.029	-0.017
	(0.029)	(0.029)	(0.029)	(0.032)	(0.029)
IT*Traditional Practices Index 3 (TPI_2)	0.020	0.020	0.008	0.007	0.010
	(0.030)	(0.029)	(0.027)	(0.027)	(0.027)
IT*Assessment		0.000	-0.002	-0.003	-0.013
		(0.019)	(0.019)	(0.019)	(0.020)
IT*Teacher demographics	.	.	√	√	√
IT*Class size	.	.	.	√	√
IT*Teacher behaviour	.	.	.	.	√

Note: This table replicates table 3, using two distinct Traditional Practices Indexes instead of one unique Modern Index. TPI_2 includes practices (a), (c) and (d) and TPI_2 includes practices (b), (e) and (f). All regressions include student and teacher fixed effects, as well as topic constants and the proportion of subtopics taught the year of the test. Controls included in columns (3) - (5) are similar to those described in table 3. Standards errors (in parentheses) are clustered at the teacher level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B15: Non linearity in the MPI effect

	Score (1)	Score (2)	Score (3)	Score (4)	Score (5)
Panel A: subtopics taught (N=18888)					
Instructional Time (IT)	0.027**	0.034***	0.017	-0.015	0.011
	(0.011)	(0.011)	(0.048)	(0.051)	(0.061)
IT*$MPI_{tophalf}$	0.032**	0.037***	0.035**	0.033**	0.027*
	(0.014)	(0.013)	(0.014)	(0.014)	(0.015)
IT*Assessment		0.044**	0.043**	0.037*	0.026
		(0.021)	(0.021)	(0.022)	(0.021)
Panel B: subtopics not taught (N=22263)					
Instructional Time (IT)	0.001	0.001	0.014	0.007	0.029
	(0.010)	(0.010)	(0.040)	(0.044)	(0.053)
IT*$MPI_{tophalf}$	-0.004	-0.004	-0.003	-0.002	-0.004
	(0.011)	(0.011)	(0.011)	(0.012)	(0.011)
IT*Assessment		-0.003	-0.003	-0.003	-0.015
		(0.018)	(0.019)	(0.019)	(0.020)
IT*Teacher demographics	.	.	√	√	√
IT*Class size	.	.	.	√	√
IT*Teacher behaviour	.	.	.	.	√

Note: This table shows the heterogeneity in the effect of math instructional time on student math performance according to the position of the teacher in the Modern Practices Index (MPI) distribution, separately on subtopics taught the year of the test (Panel A) and subtopics not taught the year of the test (Panel B). $MPI_{tophalf}$ is a dummy indicating whether the teacher ranks above the median of the MPI. All regressions include student and teacher fixed effects, as well as topic constants and the proportion of subtopics taught the year of the test. Controls included in columns (3) - (5) are similar to those described in table 3. Standards errors (in parentheses) are clustered at the teacher level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B16: Robustness check - main regressions without *Geometry*

	Score (1)	Score (2)	Score (3)	Score (4)	Score (5)
Panel A: subtopics taught (N=13899)					
Instructional Time (IT)	0.038***	0.047***	-0.016	-0.037	-0.004
	(0.010)	(0.010)	(0.039)	(0.046)	(0.078)
IT*Modern Practices Index	0.099***	0.112***	0.099***	0.102***	0.089**
	(0.036)	(0.036)	(0.035)	(0.035)	(0.039)
IT*Assessment		0.036	0.032	0.029	0.020
		(0.024)	(0.024)	(0.025)	(0.025)
Panel B: subtopics not taught (N=16711)					
Instructional Time (IT)	0.012	0.009	0.008	0.007	0.020
	(0.008)	(0.010)	(0.038)	(0.043)	(0.055)
IT*Modern Practices Index	0.014	0.013	0.017	0.018	0.013
	(0.028)	(0.028)	(0.027)	(0.028)	(0.029)
IT*Assessment		-0.012	-0.012	-0.012	-0.016
		(0.017)	(0.018)	(0.018)	(0.020)
IT*Teacher demographics	.	.	√	√	√
IT*Class size	.	.	.	√	√
IT*Teacher behaviour	.	.	.	.	√

Note: This table shows the heterogeneity in the effect of math instructional time on student math performance according to the teaching practices implemented by the math teacher, separately on subtopics taught the year of the test (Panel A) and subtopics not taught the year of the test (Panel B) and excluding *Geometry* subtopics. All regressions include student and teacher fixed effects, as well as topic constants and the proportion of subtopics taught the year of the test. Controls included in columns (3) - (5) are similar to those described in table 3. Standards errors (in parentheses) are clustered at the teacher level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B17: Instructional Productivity and Teaching Practices - Heterogeneity according to student gender

	Girls				Boys			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: topics taught this year (N_{girls}=9489; N_{boys}=9399)								
Instructional Time (IT)	0.065***	0.071***	0.014	0.026	0.032**	0.049***	-0.012	0.025
	(0.012)	(0.013)	(0.052)	(0.073)	(0.014)	(0.015)	(0.083)	(0.096)
IT*Modern Practices Index	0.118***	0.126***	0.112***	0.111**	0.068	0.094*	0.099*	0.087
	(0.040)	(0.040)	(0.042)	(0.048)	(0.057)	(0.055)	(0.057)	(0.062)
IT*Assessment		0.024	0.020	0.015		0.077**	0.067*	0.051
		(0.027)	(0.027)	(0.027)		(0.035)	(0.036)	(0.035)
Panel B: topics not taught this year (N_{girls}= 11459; N_{boys}=10804)								
Instructional Time (IT)	0.007	0.007	0.025	0.002	-0.009	-0.010	-0.024	0.052
	(0.011)	(0.012)	(0.070)	(0.077)	(0.014)	(0.016)	(0.067)	(0.102)
IT*Modern Practices Index	-0.008	-0.008	0.015	0.006	0.020	0.019	0.023	0.007
	(0.030)	(0.031)	(0.033)	(0.033)	(0.046)	(0.046)	(0.046)	(0.043)
IT*Assessment		0.003	0.003	-0.005		-0.004	-0.002	-0.021
		(0.022)	(0.022)	(0.023)		(0.032)	(0.033)	(0.036)
IT*Teacher demographics	.	.	√	√	.	.	√	√
IT*Class size	.	.	√	√	.	.	√	√
IT*Teacher behaviour	.	.	.	√	.	.	.	√

Note: This table shows the heterogeneity in the effect of math instructional time on student math performance according to the teaching practices implemented by the math teacher, separately on subtopics taught the year of the test (Panel A) and subtopics not taught the year of the test (Panel B), and separately for girls (columns (1)-(4)) and boys (columns (5)-(8)). All regressions include student and teacher fixed effects, as well as topic constants and the proportion of subtopics taught the year of the test. Controls included in columns (3) - (5) are similar to those described in table 3. Standards errors (in parentheses) are clustered at the teacher level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B18: The Modern Practices Index and Student Non Cognitive outcomes

	(1) Intrisic motivation	(2) Extrinsic motivation	(3) Self-confidence
Modern Practices Index	0.415*	0.423***	0.189
	(0.213)	(0.145)	(0.213)
Observations	7463	7459	7470

Note: This table shows the results of the regression of student non cognitive outcomes on the Modern Practices Index, controlling for student mean score in math subtopics not taught the year of the test, gender, age, socio-economic background, language spoken at home, math instructional time per week, school size, indexes of school immediate area's economic affluence and urban density and all teacher characteristics included in equation (3). Standard errors (in parentheses) are clustered at the teacher level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.