Do Teaching Practices Influence Student Achievement? *

Guido Schwerdt* Amelie C. Wuppermann†

January 30, 2009

Abstract

Recent studies have come to the conclusion that teachers matter for student learning and achievement. But up to now it remains uncertain what actually makes a good teacher. This study tries to shed light on the determinants of teacher quality by analyzing the relationship between a teaching practice and student achievement. The teaching practice looked at is the percentage of time the teacher spends in class giving lecture style presentation compared to having students solve problems. The study is based on matched student-teacher data for the United States from the 2003 Trends in International Math and Science Study (TIMSS). The data set allows to control for subject constant student traits by looking at test score differences between two subjects, math and science. While most teacher characteristics and qualifications do not explain the within student difference in test scores, more time spent on lecture style presentation is found to be an important determinant and therefore is likely an important factor related to teacher quality.

JEL-Code: I21, C21.
Keywords: Teaching practices, TIMSS, between subject variation

---

*This research has profited from comments by seminar participants at the University of Munich and the ifo Institute. We received insightful comments from Ludger Wößmann, Erik Hanushek and Jane Cooley to whom our gratitude goes.

*CESifo and Ifo Institute for Economic Research, Poschingerstr. 5, 81679 Munich, Germany. Email: schwerdt@ifo.de

†University of Munich, Kaulbachstr. 45, 80539 Munich, Germany. Email: amelie.wuppermann@lrz.uni-muenchen.de

1 Introduction

In 1966, the Coleman Report (Coleman, 1966) found that schools and teachers were not important in determining student achievement when a student's family background and his environment provided by peers and neighborhood were taken into account. Nearly 400 studies have been undertaken during the subsequent 30 years. The conclusions about the importance of schools and teachers for student achievement were diverse (Wenglinsky, 2002). With the availability of better data and methods recent studies have come to the conclusion that schools and teachers influence student learning. But the question, what actually determines teacher quality, i.e. what makes one teacher more successful in enhancing her students' performance than another, has not been settled so far (Aaronson et al., 2007).

There are three different categories of teacher variables that have been analyzed. Some studies have tried to uncover the relationship between student outcomes and teacher qualifications such as teaching certificates, other paper qualifications or teaching experience. Others look at the impact of teacher characteristics like teacher's gender and race on teacher quality. A third branch of studies focuses on the impact of teaching practices on student outcomes.[1] Teacher variables that belong to the first two categories are generally found to have only a small impact on student achievement and the part of teacher quality explained by observable teacher characteristics and certification is very small (Aaronson et al., 2007; Rivkin et al., 2005). This paper therefore tries to shed light on a question that belongs to the third branch of studies, namely, if teaching students by giving lecture style presentations instead of having them work on problems can explain differences in student achievement and teacher quality.

In accordance with the previous literature we find that many of the teacher characteristics and qualifications are not significantly related to student achievement or teacher quality when controlling for student background and all other influences that are constant across subjects. The analyzed teaching practice however is found to be significantly related to higher stu-

[1] For an overview see Goe (2007).

dent test scores. In the baseline specification one percent of teaching time spent with lecture style presentation instead of problem solving goes together with an increase of between 9 and 16 percent of a standard deviation in test scores. Even though this cannot be interpreted as a causal effect as teacher ability and personality are not accounted for it is an important result: either good teachers systematically spend more time teaching using lecture style or lecture style presentation is just a good teaching method.

The remainder of the paper is structured as follows: in the next section the literature on teaching practices is summarized. Section 3 gives an overview over the data used. Section 4 describes the estimation strategy. Baseline results are presented in Section 5, Section 6 illustrates different robustness checks and Section 7 concludes.

2 Literature on Teaching Practices

A majority of studies considering teaching practices do not look at a certain teaching practice but rather analyze the relationship between a teacher's evaluation score on a standard-based teacher evaluation system and student achievement.[2] Most of these studies find that evaluation scores are correlated with student achievement. A similar result is found by Jacob and Lefgren (2008). The authors analyze the relationship between the school principal's evaluation of a teacher and the part of actual achievement gain students have because they are taught by this teacher. The different evaluation schemes measure a part of teacher quality. Nevertheless, when analyzing the relationship between an evaluation score and student achievement it is unclear, which part of the evaluated practices is (most) important for the student outcome.

This problem also arises in some other studies that look at the impact of different categories of practices on student achievement. Smith et al.

[2] For example Borman and Kimball (2005), Gallagher (2004), Heneman et al. (2006), Holtzapple (2003), Kimball et al. (2004), Milanowski (2004), Matsumura et al. (2006) and Schacter and Thum (2004).

(2001) analyze if didactic or interactive teaching methods are more effective in teaching elementary school children. They find that interactive teaching is associated with higher gains in test scores. McGaffrey et al. (2001) and Cohen and Hill (2000) analyze if students have higher test scores in math if their teacher uses methods in accordance with a teaching reform promoted by the National Science Foundation. Again, didactic and interactive methods or reform-based and traditional practices are measured at an aggregated level encompassing different teaching practices. The authors estimate an effect of a teaching style but not of a single teaching practice.

Only a few studies have analyzed the impact of single teaching practices. Matsumura et al. (2002) look at the effect the quality of assignments has on student achievement. Using hierarchical linear modeling they find that a small part of student test score variance can be predicted by assignment quality. The relationship between assignments and student achievement is also analyzed by Newmann et al. (2001). The authors find that more intellectually challenging assignments are related to higher gains in test scores. Wenglinsky (2000, 2002) uses multilevel structural equation modeling to analyze the impact of different teaching practices on student test scores in math and science. He finds that the use of hands-on learning activities like solving real world problems and working with objects, an emphasis on thinking skills and frequent traditional testing of students, but also more individualized assessment through projects and portfolios are positively related to students' test scores taking into account student background and prior performance. Some evidence for the effectiveness of frequent student assessment is also found by Kannapel et al. (2005): High-performing high-poverty schools in Kentucky payed more attention to student assessment than other high-poverty schools. Bonesrønning (2004) looks at a different aspect of student assessment. He analyzes if grading practices affect student achievement in Norway and finds evidence that easy grading deteriorates student achievement.

Brewer and Goldhaber (1997) estimate different specifications of education production functions for tenth grade students in math with data from the National Educational Longitudinal Study of 1988. They conclude that teacher behavior is important in explaining student test scores. Especially, they find that controlling for student background, prior performance and school and teacher characteristics, instruction in small groups and emphasis on problem solving lead to lower student test scores.

Similarly to Brewer and Goldhaber (1997), this paper analyzes the effect of single teaching practices on student achievement. But the analyzed practice is different from those of Brewer and Goldhaber (1997). As in Brewer and Goldhaber (1997) problem solving is included in the analysis. But it is not taken as the mainly analyzed teaching practice. Instead we look at the effect of spending time on lecture style presentation compared to time spend on problem solving. Since lecture style presentation and problem solving could be classified as belonging to different teaching styles this study also relates to other literature that compares the effects of different teaching styles like Smith et al. (2001).

3 Data

The data used here is the 2003 wave of the Trends in International Math and Science Study (TIMSS). About 50 countries participated in the study. As most of the studies on the effect of teaching practices concern US data this analysis also looks at students from the US. In TIMSS, students in 4th grade and in 8th grade were tested in math and science. Test scores were standardized across countries to a mean of 500 and a standard deviation of 100. For the following analysis, we standardize the test scores for each subject to be mean 0 and standard deviation 1.[3] In addition to test scores, the publicly available TIMSS data set provides background information on student home and family. For the purpose of estimating teacher effects on student outcomes it is very important that TIMSS is a matched student-teacher data set. Each student's teachers in math and science were surveyed on their characteristics, qualifications and teaching practices. Questions concerning teaching practices were also answered by the students. Data on certain school

[3]Different approaches were taken for standardization that all took the sampling design into account. In a first approach I standardized the test scores for the whole US sample. In a second approach test scores were standardized within the actually used sample. Results do not differ for the two approaches.

characteristics reported by each school's principal is also available.

Besides the matched student-teacher design of the data it is important to have test scores for each student in math and science. The availability of test scores in two subjects allows to control for all observed and unobserved student traits that are constant across the two subjects and have the same influence on both subjects test scores (Dee, 2005, 2007).

The teaching practice analyzed in this paper is derived from question 20 in the teacher questionnaires. It asks teachers to report what percentage of time in a typical week of the specific subject's lessons students spend on reviewing homework, listening to lecture style presentation, working on problems with the teacher's guidance, working on problems without guidance, listening to the teacher reteach and clarify content, taking tests or quizzes, classroom management and other activities. Out of these 8 categories, we classify listening to lecture style presentation and working on problems with and without guidance as effective teaching time, that is time in which students study new material. The percentage of time spent on effective teaching is included as a control in the following analysis. As we are especially interested in the impact of more traditional teaching methods like giving lecture style presentations compared to more modern and interactive methods like having students solve problems in class we generate the share of time spent on lecture style presentation on the overall effective teaching time. That is the analysis focuses on the relationship between the amount of time spent on lecture style presentation relative to time spent on problem solving.

The TIMSS 2003 US data set contains student-teacher observations on 8.912 students in 232 schools. 41 of those students have more than one teacher in science. These students are dropped from the analysis. 8.871 students in 231 schools in 455 math classes taught by 375 different math teachers and in 1.085 science classes taught by 475 different science teachers remain in the sample. Not all of the students and teachers completed their questionnaires. In order not to be forced to drop a large amount of observations we included dummy variables that indicate if a certain categorial variable is missing. 2.605 students are dropped nevertheless due to missing teaching practice variables. As the test scores of the dropped students are a little lower on average than the average test scores in the remaining sample the results have to be interpreted with caution. 6.305 students in 204 schools with 638 teachers (284 math teachers and 335 science teachers, and 19 teachers who teach both subjects) remain in the sample.

Another feature of the data is that many students have different peers in math and science. The TIMSS data permits to identify class and peers for each student in both subjects. This allows to construct average class variables which - as a robustness check - will be used as instruments for the teacher responses on teaching practices. Also, as a robustness check, peer effects will be eliminated by restricting the sample to students who have the same peers in math and science.

Furthermore due to the sampling design of TIMSS, students are not all selected with the same probability. A two stage sampling design makes it necessary to take probability weights into account when estimating summary statistics (Martin, 2005; Cameron and Trivedi, 2005). This is done for the statistics shown in Table 1. The results of the regression analyses reported below also take the probability weights into account and allow for correlation between error terms within schools.[4]

4 Estimation Strategy

We estimate a standard education production function using OLS and a first difference approach which helps to control for student ability.

In an education production function the test score (Y_{ij}) of student i in subject j is explained by student background characteristics B_i, school characteristics S_i, peer characteristics P_{ij} and teacher characteristics and qualifications T_{ij}. In addition, teaching practices TP_{ij} are included in the following analysis.

$$Y_{ij} = c_j + B_i'\beta_{1j} + S_i'\beta_{2j} + P_{ij}'\beta_{3j} + T_{ij}'\beta_{4j} + TP_{ij}'\beta_{5j} + \epsilon_{ij} \tag{1}$$

[4]In addition, the two step procedure of sampling could be incorporated in the estimation of standard errors. For simplicity, we ignore the latter in the following analysis which then gives us conservative estimates of the standard errors StataCorp (2007).

where ϵ_{ij} contains all unobservable influences on student test scores. In particular, it contains the effects of unobservable student traits, like student ability, and unobservable teacher, school and peer characteristics. OLS results might therefore be biased due to omitted variables. The results merely represent conditional correlations of teaching practices and test scores.

All constant student traits that have the same influence on the two subjects can be controlled for by differencing the equations for the two subjects. This approach closely follows Dee (2005, 2007). Denoting the difference between variables in two subjects as ΔVar, e.g. the difference in test scores as $\Delta Y_i = Y_{im} - Y_{is}$, the equation becomes:

$$\begin{aligned}\Delta Y_i &= c_m - c_s + B_i'(\beta_{1m} - \beta_{1s}) + S_i'(\beta_{2m} - \beta_{2s}) \\ &\quad + P_{im}'\beta_{3m} - P_{is}'\beta_{3s} + T_{im}'\beta_{4m} - T_{is}'\beta_{4s} + TP_{im}'\beta_{5m} - TP_{is}'\beta_{5s} + \eta_i \end{aligned} \tag{2}$$

where $\eta_i = \epsilon_{im} - \epsilon_{is}$. Everything that is included in both subjects' error terms ϵ_{ij} thus cancels out. In a baseline specification, we follow a traditional first difference approach assuming that the coefficients for each variable are equal across the two subjects. This way all variables that are constant across subjects are differenced out and the equation becomes:

$$\Delta Y_i = \Delta T_i'\beta_4 + \Delta TP_i'\beta_5 + \eta_i \tag{3}$$

As a robustness check we also estimate equation (2) which especially allows subject constant variables to have differential effects in the two subjects.

To estimate a causal effect of a teaching method we have to assume that no unobserved student, teacher, school or peer characteristics contained in the error term η_i are correlated with the explanatory variables.

But even in the first differenced equation it is likely that the teaching practices are endogenous. On the one hand, teachers choose the practice according to the specific situation in each classroom which itself might determine student achievement. On the other hand, the effect of teaching methods might be related to the personality and ability of each teacher. To see how important the first concern is, further teacher variables are included. These variables come from the teacher questionnaires and report how much each teacher feels hindered from teaching by unmotivated students or missing resources.[5] To find further evidence for our finding we will try to find instruments that provide exogenous variation in teaching practices.

5 Results

Equation (1) is estimated on a pooled sample which includes each student twice, once with the math test score and math teacher variables and once with the science test score and science teacher variables. This approach assumes that the coefficients β_j are the same for the two subjects. In addition, Table 2 reports estimates for the two subjects separately, allowing β_j to vary between the subjects.

The results for the pooled OLS estimation reported in Table 2 show that student variables are strongly related to student achievement. Boys do better than girls, students with a migration background do worse than students whose parents have already been born in the US and older students do worse than younger students. These results are stable across subjects, and also hold when controlling for across school sorting by including school fixed effects.

The coefficients of the teacher variables are hardly significantly different from 0. Only a teacher's gender seems to be significantly related to student test scores: Female teachers have students with significantly lower test scores in both subjects.

The relationship between more lecture style presentation and test scores is positive and significantly different from zero in all but one specification. Controlling for between school sorting does not change this relationship. That is teachers who spend relatively more time in class on lecture style presentation teach students who have higher test scores in the TIMSS study. These results might nevertheless just be driven by student ability: teachers who have higher ability students can teach more abstract classes and thus spend more time on lecturing than on problem solving.

[5]These variables are referred to as "Teaching Limits" in the following text.

Looking at the results of the first difference estimation in Table 3 student ability does not seem to be the driving factor behind the relationship between teaching method and test scores. When regressing the difference of each student's test scores in math and science on the difference of teacher reports between the subjects the estimate of the coefficient for the percentage of effective teaching time spent on giving lecture style presentations is again significantly positive. So, if one of the two teachers reports a higher share of giving lecture style presentations in his effective teaching time the student does significantly better in his subject than in the other teacher's subject. The magnitude of the effect also seems to be quite impressive: one percent more spent on lecture style presentation instead of problem solving is related to an increase in test scores of between .09 and .16, which is equal to 9 and 16 percent respectively of a standard deviation in test scores. Interestingly, hardly any of the other teacher variables have a significant relationship to the difference in test scores. It does not seem to matter for the student's achievement whether his teacher has a major in a math or science, whether his teacher is female or male, whether his teacher is old or young and whether his teacher has participated in a lot of teacher training activities or not. Only a major in education and teaching experience have a significant coefficient in one of the specifications. The latter effect nicely aligns with results in other studies that find that teaching experience is one of the few variables that have a significant impact on teacher quality (Rivkin et al 2005).

As pointed out when describing the estimation strategy the first difference approach takes out all influences on student test scores that are constant across the subjects. But teacher differences in ability or personality remain in the error term and likely result in inconsistent estimates. This problem can be mitigated with an instrumental variable approach, provided that we are able to find convincing instruments.

All in all, time in class spent on giving lecture style presentation seems to be an important factor in influencing student achievement. How sensitive the results are to changes in the specification and scaling of the teaching method variable is described in the next section.

6 Robustness Checks

Table 4 reports different robustness checks to the results presented thus far. In columns (2) and (3) we allow the coefficients of the different teacher variables to vary between the two subjects by estimating equation (2). In addition, column (1) reports the results of additionally controlling for more of the other categories in question 20 in the teacher questionnaire and column (4) instruments teacher responses with student responses.

The results in columns (2) and (3) show that the positive effect of teaching with lecture style presentation might be more pronounced in math than in science. The coefficients for both subjects show the expected sign. Given that we abstract from spillover effects between the subjects, that is the teaching in math is assumed to have no impact on the test scores in science and vice versa, a positive estimate for the math coefficient can be interpreted as a positive relationship between teaching with lecture style in math and math test scores. As the science test scores enter with a negative sign in the dependent variable a negative coefficient for lecture style presentation in science mirrors a positive relationship of more lecture style presentation and test scores in science. The latter, however, are not significant in this specification.

This specification also allows to include student background variables as additional controls as shown in column (3) of Table 4. As the differences between the test scores in the two subjects are taken to control for student ability, it is reassuring, that the results do not change when including student background variables. Even though student's age, gender and migration background explain part of the variation in the test score difference, the relationship between teaching practice and the test score difference is not affected by the inclusion of background variables.

As the results in column (1) of Table 4 show controlling for the other categories in question 20 of the teacher questionnaire does not change the coefficient of lecture style presentation. It is still positive and significantly different from zero. The other categories itself also do not seem to be related significantly to student test scores.

So far, results have been shown using the teacher self-reported teaching practice. In column (4) of Table 4 student reported teaching practices are used to instrument teacher responses. When using instrumental variables the coefficient of lecture style presentation is no longer significant.

Except for the results in the instrumental variables estimation the robustness checks reported in this section further support the relationship found between relatively more time spent on giving lecture style presentation in class and student test scores. Further robustness checks as for example controlling for peer effects by looking only on students who have the same peers in the two subjects and only at students in schools in which students are not sorted into different classes according to ability ("‘no tracking schools"’) are to follow.

7 Conclusion

Overall, spending more time in class on lecture style presentation than on problem solving is found to be related to higher student achievement and thus higher quality teachers. This is, to some extent, contrary to the result of Smith et al. (2001), who find that interactive teaching enhances student performance to a greater extent than the use of didactic methods. But these authors analyze categories of teaching practices and not single practices and look at children in elementary school whereas we focus on students in 8th grade. It might be that different teaching practices are appropriate for different ages.

It is important to repeat that the estimated relationships should not be interpreted as causal effects. The OLS and first difference approaches both likely suffer from omitted teacher variables. It could be that good teachers spend more time on lecture style presentation, but that they would be equally successful if they had their students solve problems more often. Nevertheless, it could also be that teaching with lecture style presentation really is a more successful method than problem solving. The results of the instrumental variable estimation will hopefully shed light on this question.

As a further line of research we want to analyze if the relationship between lecture style presentation and test scores depends on the level of lecture style presentation. In addition, the analysis could be expanded to 4th grade students in the US and/or students in earlier waves of TIMSS to see how robust the relationships are.

References

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

Bonesrønning, H. (2004). Do the teachers' grading practices affect student achievement? *Education Economics*, 12(2):151–167.

Borman, G. D. and Kimball, S. M. (2005). Teacher quality and educational quality: Do teachers with higher standards-based evaluation ratings close student achievement gaps? *The Elementary School Journal*, 106(1):3–20.

Brewer, D. J. and Goldhaber, D. (1997). Why don't schools and teachers seem to matter? *The Journal of Human Resources*, 32(3):505–523.

Cameron, A. C. and Trivedi, P. K. (2005). *Microeconometrics*. Cambridge University Press.

Cohen, D. K. and Hill, H. C. (2000). Instructional policy and classroom performance: The mathematics reform in california. *Teachers College Record*, 102(2):294–343.

Coleman, J. S. (1966). Equality of educational opportunity. Technical report, Washington DC: US Government Printing Office.

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *American Economic Review*, 95(2):158–165.

Dee, T. S. (2007). Teachers and the gender gaps in student achievement. *The Journal of Human Resources*, 42(3):528–554.

Gallagher, H. A. (2004). Vaughn elementary's innovative teacher evaluation system: Are teacher evaluation scores related to growth in student achievement? *Peabody Journal of Education*, 74(4):79–107.

Goe, L. (2007). The link between teacher quality and student outcomes: A research synthesis. Report, National Comprehensive Center for Teacher Quality, Washington DC.

Heneman, H. G., Milanowski, A., Kimball, S. M., and Odden, A. (2006). Standards-based teacher evaluation as a foundation for knowledge- and skill-based pay. CPRE Policy Brief RB-45, Consortium for Policy Research in Education, Philadelphia.

Holtzapple, E. (2003). Criterion-related validity evidence for standard-based teacher evaluation systems. *Journal of Personnel Evaluation in Education*, 17(3):207–219.

Jacob, B. A. and Lefgren, L. (2008). Can principals identify effective teachers? evidence on subjective performance evaluation in education. *Journal of Labor Economics*, 26:101–136.

Kannapel, P. J., Clements, S. K., Taylor, D., and Hibpshman, T. (2005). Inside the black box of high-performing high-poverty schools. Technical report, Prichard Committee for Academic Excellence.

Kimball, S. M., White, B., Milanowski, A. T., and Borman, G. (2004). Examining the relationship between teacher evaluation and student assessment results in washoe county. *Peabody Journal of Education*, 79(4):54–78.

Martin, M. E. (2005). *TIMSS 2003 User Guide for the International Data Base*. International Association for the Evaluation of Educational Achievement (IEA), Boston.

Matsumura, L. C., Garnier, H. E., Pascal, J., and Valdes, R. (2002). Measuring instructional quality in accountability systems: Classroom assignments and student achievement. *Educational Assessment*, 8(3):207–229.

Matsumura, L. C., Slater, S. C., Junker, B., Peterson, M., Boston, M., Steele, M., and Resnick, L. (2006). Measuring reading comprehension and mathematics instruction in urban middle schools: A pilot study of the instructional quality assessment. CSE Technical Report 681, National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles.

McGaffrey, D. F., Hamilton, L. S., Stecher, B. M., Klein, S. P., Bugliari, D., and Robyn, A. (2001). Interactions among instructional practices, curriculum and student achievement: The case of standards-based high school mathematics. *Journal for Research in Mathematics Education*, 35(5):493–517.

Milanowski, A. (2004). The relationship between teacher performance evaluation scores and student achievement: Evidence from cincinnati. *Peabody Journal of Education*, 79(4):33–53.

Newmann, F. M., Bryk, A. S., and Nagaoka, J. (2001). Authentic intellectual work and standardized tests: Conflict or coexistence? Technical report, Consortium on Chicago School Research, Chicago.

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

Schacter, J. and Thum, Y. M. (2004). Paying for high- and low-quality teaching. *Economics of Education Review*, 23(4):411–430.

Smith, J. B., Lee, V. E., and Newmann, F. M. (2001). Instruction and achievement in chicago elementary schools. Report, Consortium on Chicago School Research, Chicago.

StataCorp (2007). *Stata Survey Data Reference Manual*. StataCorp LP, Texas, 10 edition.

Wenglinsky, H. (2000). How teaching matters: Bringing the classroom back into discussion of teacher quality. Technical Report Princeton, Policy Information Center.

Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and student academic performance. *Education Policy Analysis Archives*, 10(12).

Table 1: Descriptive Statistics

Variable	Mean[1]	SD[1]	Min	Max	N	
Student Background					6305	
Age	14.226	.463	12.50	17.17		
Firsthalf	.498	.500	0	1		Dummy if born in first 6 months
Female	.522	.500	0	1		Dummy if female
Books2	.174	.379	0	1		Dummy if 11-25 books at home
Books3	.279	.448	0	1		Dummy if 26-100 books at home
Books4	.180	.384	0	1		Dummy if 101-200 books at home
Books5	.249	.433	0	1		Dummy if >200 books at home
Language	2.769	.594	0	3		Frequency of English spoken at home - 0 never to 3 always
Immigrant	.144	.351	0	1		Dummy if migration background
no. people	4.504	1.338	2	8		Number of people at home
Educpar	3.126	1.135	0	4		Parental Education - no more than 0 primary to 4 finished university
Teacher Variables						
Math T[2]					303	Math Teacher
Math T mathmaj	.458	.499	0	1	288	Dummy if major in math
Math T sciencemaj	.142	.350	0	1	288	Dummy if major in science
Math T educmaj	.577	.495	0	1	291	Dummy if major in education
Math T Age	3.747	1.142	1	6	297	Age categories - 1 under 25 to 60 or older
Math T Female	.639	.481	0	1	296	Dummy if female
Math T Exp	14.348	10.193	1	47	293	Years of teaching experience
Math T training	1.451	1.1204	0	6	297	Years of teacher training
Sci T[2]					354	Science Teacher
Sci T mathmaj	.093	.290	0	1	335	Dummy if major in math
Sci T sciencemaj	.615	.487	0	1	343	Dummy if major in science
Sci T educmaj	.454	.499	0	1	339	Dummy if major in education
Sci T Age	3.754	1.157	1	6	346	Age categories - 1 under 25 to 60 or older
Sci T Female	.545	.499	0	1	347	Dummy if female
Sci T Exp	14.090	10.694	1	50	343	Years of teaching experience
Sci T training	1.441	1.199	0	6	347	Years of teacher training

[1] *Probability weights and within school correlation were taken into account when estimating means and standard deviations.*
[2] *The teacher variables are shown as subject specific teacher averages. Subject specific student averages of teacher variables are very similar.*

Table 2: Estimation Results OLS

	Math		Science		Pooled	
Lecture style teaching	.601***	.645*	.107	.693**	.343***	.284***
	(.20)	(.38)	(.12)	(.29)	(.11)	(.09)
Effective teaching time	.210	1.024**	.0248	.439	.0795	.347***
	(.29)	(.45)	(.15)	(.28)	(.14)	(.13)
Age	−.186***	−.181***	−.109***	−.101***	−.147***	−.143***
	(.03)	(.02)	(.03)	(.02)	(.03)	(.02)
Born in first 6 months	−.0144	−.0152	−.0253	−.0361*	−.0194	−.0247
	(.02)	(.02)	(.02)	(.02)	(.02)	(.02)
Female	−.131***	−.121***	−.278***	−.265***	−.205***	−.193***
	(.02)	(.02)	(.02)	(.02)	(.02)	(.02)
Immigrant	−.220***	−.155***	−.318***	−.268***	−.273***	−.213***
	(.05)	(.04)	(.05)	(.04)	(.04)	(.04)
Teacher's gender	−.121*	−.0814	−.0795	−.0224	−.0909*	−.0303
	(.06)	(.13)	(.06)	(.09)	(.05)	(.04)
Teacher's age	−.0326	.0386	−.0156	−.0921*	−.0272	−.0275
	(.04)	(.08)	(.03)	(.05)	(.03)	(.03)
Teacher's experience	.00671	.00762	.00138	.00361	.00418	.00337
	(.00)	(.01)	(.00)	(.01)	(.00)	(.00)
Teacher's training	−.00854	.0227	−.0321	−.0134	−.0199	−.00970
	(.03)	(.07)	(.02)	(.03)	(.02)	(.01)
Teacher's major math	.0913	.0101	.123	.156	.0584	.00121
	(.06)	(.14)	(.09)	(.16)	(.05)	(.04)
Teacher's major science	−.157*	−.0668	.00295	−.123	−.0182	.00252
	(.09)	(.20)	(.06)	(.08)	(.04)	(.03)
Teacher's major education	.0579	.0960	.0733	.00990	.0603	.0460
	(.07)	(.14)	(.06)	(.09)	(.04)	(.04)
Missing Categories	Yes	Yes	Yes	Yes	Yes	Yes
Family background	Yes	Yes	Yes	Yes	Yes	Yes
School characteristics	Yes	No	Yes	No	Yes	No
School fixed effects	No	Yes	No	Yes	No	Yes
Constant	−.311	−1.375**	−.0691	−1.350**	−.575	−.128
	(.61)	(.68)	(.48)	(.55)	(.51)	(.39)
Observations	6305	6305	6305	6305	12610	12610
R^2	.292	.494	.335	.501	.303	.476

Note: Clustered standard errors in parenthesis.

Table 3: Estimation Results First Difference

	1	2	3	4
Lecture style teaching	.0923*	.0922*	.0995**	.157*
	(.05)	(.05)	(.05)	(.09)
Effective teaching time		−.00545	.0239	.112
		(.06)	(.06)	(.10)
Diff. teachers' gender			−.0170	−.0425
			(.02)	(.03)
Diff. teachers' age			.0206*	−.01000
			(.01)	(.02)
Diff. teachers' experience			−.0000204	.00386*
			(.00)	(.00)
Diff. teachers' training			−.00327	−.0203
			(.01)	(.01)
Diff. major math			.00151	−.0633
			(.02)	(.05)
Diff. major science			−.0107	.0369
			(.02)	(.03)
Diff. major education			.0162	−.102***
			(.02)	(.03)
Missing Categories	Yes	Yes	Yes	Yes
Limit to teach	Yes	Yes	Yes	Yes
School fixed effects	No	No	No	Yes
Constant	−.0104	−.0103	−.0128	.123
	(.02)	(.02)	(.02)	(.11)
Observations	6305	6305	6305	6305
R^2	.020	.020	.025	.095

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

Note: Clustered standard errors in parenthesis.

Table 4: Estimation Results Robustness

	1	2	3	4
Lecture style teaching	.118**	.141*	.143*	.188
	(.06)	(.08)	(.08)	(.98)
Lecture style teaching in science		−.0847	−.0920	
		(.07)	(.07)	
Effective teaching time in science		−.0478	−.0257	
		(.08)	(.07)	
Effective teaching time	−.911	−.0244	−.0281	
	(1.03)	(.11)	(.11)	
Teaching time	−1.143			
	(1.07)			
Not other stuff	−.782			
	(.69)			
Diff. in Teachers' characteristics	Yes	No	No	Yes
Teachers' characteristics	No	Yes	Yes	No
Science Teachers' characteristics	No	Yes	Yes	No
Personal characteristics	No	No	Yes	No
Family background	No	No	Yes	No
Observations	6305	6305	6305	6305
R^2	.006	.008	.037	.

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

Note: Clustered standard errors in parenthesis.