Program on Education Policy and Governance Working Papers Series

Is traditional teaching really all that bad?
A within-student between-subject approach

Guido Schwerdt
CESifo and Ifo Institute for
Economic Research

Amelie C. Wuppermann
University of Munich

PEPG 10-15

Program on Education Policy and Governance Harvard Kennedy School 79 JFK Street,
Taubman 304 Cambridge, MA 02138 Tel: 617-495-7976 Fax: 617-496-4428
www.hks.harvard.edu/pepg/

Is traditional teaching really all that bad? A within-student between-subject approach

Guido Schwerdt* Amelie C. Wuppermann†

June 19, 2010

Abstract

Recent studies conclude that teachers are important for student learning but it remains uncertain what actually determines effective teaching. This study directly peers into the black box of educational production by investigating the relationship between lecture style teaching and student achievement. Based on matched student-teacher data for the US, the estimation strategy exploits between-subject variation to control for unobserved student traits. Results indicate that traditional lecture style teaching is associated with significantly higher student achievement. No support for detrimental effects of lecture style teaching can be found even when evaluating possible selection biases due to unobservable teacher characteristics.

JEL-Code: I21, C23.
Keywords: Teaching practices, Educational production, TIMSS, Between-subject variation

This research has profited from comments by seminar participants at the EALE meeting in Tallinn, the EEA meeting in Barcelona, the EDGE meeting in Copenhagen, the University of Munich and the Ifo Institute. We received insightful comments from Ludger Woessmann, Erik Hanushek, Joachim Winter, Florian Heiß and Jane Cooley to whom our gratitude goes. Financial support from the Deutsche Forschungsgemeinschaft through GRK 801 is gratefully acknowledged.

*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

Recent studies stress the importance of teachers for student learning. However, 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). Different categories of teacher variables have been analyzed. Some studies focus on the impact of a teacher's gender and race on teacher quality (Dee, 2005, 2007). Others try to uncover the relationship between student outcomes and teacher qualifications such as teaching certificates, other paper qualifications or teaching experience (Kane et al., 2008). Such observable teacher characteristics are, however, generally found to have only little impact on student achievement and can only explain a relatively small part of overall teacher quality (Aaronson et al., 2007; Rivkin et al., 2005). Most of the variation in teacher quality can be attributed to unobserved factors.[1]

While most of these studies focus on characteristics of the teacher, this paper directly peers into the black box of educational production by focusing on the actual teaching process. More specifically, we contrast two teaching practices, teaching by giving lecture style presentations with teaching based on in-class problem solving, and investigate the impact on student achievement. Giving lecture style presentations is often regarded as old-fashioned and connected with many disadvantages: Lectures fail to provide instructors with feedback about student learning and rest on the presumption that all students learn at the same pace. Moreover, students' attention wanes quickly during lectures and information tends to be forgotten quickly when students are passive. Finally, lectures emphasize learning by listening, which is a disadvantage for students who prefer other learning styles. Alternative instructional practices based on active and problem-oriented learning presumably do not suffer from these disadvantages. National standards (NCTM, 1991; National Research Council, 1996) consequently advocate engaging students more in hands-on learning activities and group work. Despite these recommendations traditional lecture and textbook methodologies continue to dominate science and mathematics instruction in US middle schools (Weiss, 1997). This raises the question whether student achievement could be raised by reducing the high share of teaching time devoted to lecture style presentations.

By addressing this question, this study adds to the literature analyzing the impact of teaching process variables such as teaching practices on student outcomes.[2] Despite

[1]This finding led researchers, concerned with providing recommendations for recruitment policies and designing optimal teacher pay schemes, to suggest to identify effective teachers by their actual performance on the job using "value added" measures of student achievement (Gordon et al., 2006).

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

the importance of teaching practices for student performance as recognized by educational researchers (Seidel and Shavelson, 2007) and their potential relatively low-cost implementation, economists have only recently begun to analyze the impact of teaching methods on student achievement.[3] Various dimensions of teaching practices have been shown to be able to explain a large share of the between-teacher variation in student achievement (Schacter and Thum, 2004). However, to our knowledge no rigorous empirical analysis of the impact of lecture style teaching on overall student achievement exists.

To study the effect of lecture style teaching relative to teaching based on in-class problem solving we use information on in-class time use provided by teachers in the 2003 wave of the Trends in International Mathematics and Science Study (TIMSS) in US schools. Estimating a reduced form educational production function and exploiting between-subject variation to control for unobserved student traits, we find that the choice of teaching practices matters for student achievement. We find that a 10 percentage point shift from problem solving to lecture style presentation results in an increase in student achievement of about 1 percent of a standard deviation.

This result is highly robust. Consistent with other studies in this literature, we find no evidence for significant effects of commonly investigated observable teacher characteristics such as teaching certificates or teaching experience. While we are able to control for a huge array of observable teacher traits, selection of teachers based on unobservable characteristics into teaching methods remains an issue. The bias resulting from potential selection of teachers with different unobservable attributes into different teaching methods is assessed following the technique pioneered in Altonji et al. (2005). The results indicate that only relatively low selection on unobservables compared to the selection on observables is necessary to explain the entire estimated effect. We would thus not formulate policy conclusions that call for more lecture style teaching compared to problem solving in general. However, a negative causal effect of giving lecture style presentations that is hidden in our results due to selection based on unobserved teacher traits is also not very likely. It can only exist if "good" teachers (teachers with favorable unobserved characteristics) predominately select themselves into an inferior teaching technique. This scenario, however, lacks any intuitive or theoretical support and thus appears extremely implausible. We therefore conclude that the high share of total teaching time devoted to traditional lecture style teaching in

[3]Brewer and Goldhaber (1997) and Aslam and Kingdon (2007) analyze the impact of many different teaching methods. Rouse and Krueger (2004), Banerjee et al. (2007), and Barrow et al. (2009) investigate the effectiveness of computer-aided instruction and Machin and McNally (2008) analyze an education policy that changed reading instruction.

science and mathematics instruction in US middle schools has no detrimental effect on student achievement. Our findings imply that simply changing the teaching method from lecture style presentation to problem solving without concern for how the methods are implemented has little potential for raising overall achievement levels.

The remainder of the paper is structured as follows: the following section reviews the literature on teaching practices. Section 3 presents the data. Section 4 describes the estimation strategy. Headline results are presented in Section 5, while Section 6 provides a sensitivity analysis. Section 7 concludes.

2 Literature on Teaching Practices

There are two strands of literature that are closely related to our study. The first strand analyzes the impact of different teaching styles on student achievement. In these studies teaching style variables are meant to proxy for broader pedagogical concepts. A certain teaching style may consist of a combination of different teaching practices, where the term teaching practice indicates the actual classroom activity, like giving a lecture style presentation or reviewing homework. The effects of single teaching practices are analyzed by the second strand of the related literature.

Findings of the first strand of literature on teaching style speak in favor of modern, interactive teaching styles. Smith et al. (2001) find that primary school students in Chicago had higher test score gains when they were taught with an interactive teaching style compared to didactic or review-oriented teaching styles. Similarly, McGaffrey et al. (2001) find that a teaching style in accordance with a reform promoted by the National Science Foundation during the 1990s can be beneficial for student achievement in math in 10th grade. Their finding suggests that a change from traditional teaching to the reform-based style only enhances student achievement if the curriculum is changed in accordance with the reform in addition to changing the teaching style.

Another study that analyzes a simultaneous change in both, structure and content of teaching, is Machin and McNally (2008). This study analyzes the effect of the introduction of the "literacy hour" in English primary schools in the late 1990s. This policy intervention by the British government changed how primary school students are taught to read. Using the fact that not all schools started the literacy hour at the same point in time in a difference-in-difference framework, the authors show that the literacy hour significantly increased reading skills for low ability student while high ability students were not affected.

According to the findings of these studies the overall teaching style that teachers use in class seems to matter for student achievement. As these styles are, however, composed of different individual teaching practices it is hard to come up with recommendations to teachers what exactly they should do in the classroom.

The second strand of literature analyzes individual teaching practices. A first group of studies in this area analyzes the use of computer based instruction. The evidence on the use of computer programs is mixed for different subjects. On the one hand, Rouse and Krueger (2004) find that literacy skills of 3rd and 6th grade students in a large US urban school district do not profit from the use of computer programs for teaching language and reading. Students' math skills, on the other hand, are shown to increase due to computer-aided instruction as shown by Banerjee et al. (2007) for India and Barrow et al. (2009) for the US. Computer-based instruction might thus be more effective in some than in other subjects.

Other teaching practices that have received some attention in the literature are accountability measures. Wenglinsky (2000, 2002) finds that frequent testing of students is positively related to students' test scores taking into account student background and prior performance. Additional evidence for the importance of testing is presented in Kannapel et al. (2005): High-performing high-poverty schools in Kentucky payed more attention to student assessment than other high-poverty schools. In addition, there is some evidence that not only testing itself is important but also how the tests are graded. Bonesrønning (2004) analyzes if grading practices affect student achievement in Norway and finds evidence that easy grading deteriorates student achievement. The design of tests might also matter. This hypothesis is supported by findings of Newmann et al. (2001) and Matsumura et al. (2002). Both studies find evidence for a positive link between the quality of assignments that students are asked to do and overall student performance. These findings emphasize the importance of taking into account other categories of in-class time use such as accountability measures when studying the effect of teaching by giving lecture style presentations in comparison to in-class problem solving.

More closely related to this paper in terms of the teaching practices analyzed and identification strategy are the analyses by Brewer and Goldhaber (1997) and Aslam and Kingdon (2007). 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. Controlling for student background, prior performance and school and teacher characteristics, they find that instruction in small groups and emphasis on problem solving lead to lower student test scores.

Aslam and Kingdon (2007) analyze the impact of different teaching practices on student achievement in Pakistan. Their identification strategy rests on within pupil across subject (rather than across time) variation, which is similar to the identification strategy employed in this analysis. They find that students taught by teachers who spend more time on lesson planning and by teachers who ask more questions in class have higher test scores.

Similar to this second strand of studies we focus on a comparison of single teaching practices, namely time devoted to teaching by giving lecture style presentations versus time devoted to teaching by problem solving. As lecture style teaching is a teaching practice that is often associated with more traditional, didactic or teacher-centered teaching styles, while problem solving is connected to more modern, interactive or student-centered approaches to teaching, our results might be also of a more general interest. Strictly speaking, however, our data only allows to draw conclusions about actual classroom activities.

3 Data

The data used in this study is the 2003 wave of the Trends in International Mathematics and Science Study (TIMSS). In this study we focus on country information for the US. In TIMSS, students in 4th grade and in 8th grade were tested in math and science. We limit our analysis to 8th grade students, because 4th grade students are typically taught by one teacher in all subjects.

We standardize the test scores for each subject to be mean 0 and standard deviation 1. In addition to test scores in the two tested subjects, the TIMSS data provide background information on student home and family. For the purpose of this analysis it is crucial that TIMSS allows linking students to teachers. Each student's teachers in math and science are surveyed on their characteristics, qualifications and teaching practices. Additionally, school principals provide information on school characteristics.

The key variable of interest in this paper is derived from question 20 in the teacher questionnaires in the 2003 wave of TIMSS. Unfortunately, this precise question was not asked in previous waves of TIMSS. We therefore limit our analysis to the 2003 wave. Teachers are asked in 2003 to report what percentage of time in a typical week of the specific subject's lessons students spend on eight in-class activities. Our three main categories of interest are listening to lecture style presentation, working on problems with the teacher's guidance and working on problems without guidance. The overall time in class spent on these three activities likely provides a good proxy for the time in class in which students are taught new material. The main interest of this study is to contrast teaching by giving lecture style presentations to teaching based on problem solving. The two problem solving categories are thus combined into a single teaching practice: teaching based on problem solving. To allow a comparison of the two categories we construct a single variable, *lecture style teaching*, that is the percent of time spent giving lecture style presentation relative to the percent of time spent on either of the two activities ($\frac{lecture(\%)}{lecture(\%)+problemsolving(\%)}$). The key advantage is that a change in this variable can directly be interpreted as a shift from spending time on one to spending time on the other practice holding constant the overall time spent on these two practices. For example, an increase in this variable of 0.1 indicates that 10 percentage points of total time devoted to teaching new material are shifted from teaching based on problem solving to giving lecture style presentations.

The remaining 5 activities include reviewing homework, listening to the teacher reteach and clarify content, taking tests or quizzes, classroom management and other activities. We simply group these activities and construct a single variable, *other class activities*, which measures the share of total time in class spent on other activities. We include this variable as a control in all specifications. To analyze the robustness of our results we also estimate specifications that include the individual shares of all categories. Moreover, teachers also report the total time in minutes per week that they teach math or science to the class, which we also include as a control variable in most specifications.

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 not included in the estimation sample. 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 lose a large amount of observations we impute missing values of all control variables and include indicators for imputed values in all estimations.[4] 2,561 students have, however, missing information on our teaching practice variable of interest. These observation are dropped from the analysis. 6,310 students in 205 schools with 639 teachers (303 math teachers and 355 science teachers, where 19 teachers teach both subjects) remain in the sample.

Furthermore due to the sampling design of TIMSS, students are not all selected with

[4]Experimenting with different imputation procedures revealed that our main results do not depend on the method of imputation. Main results remain also qualitatively unchanged when simply deleting observations with missing values. Results presented in the paper are based on a simple mean-imputation procedure.

the same probability. A two stage sampling design makes it necessary to take probability weights into account when estimating summary statistics (Martin, 2005). All estimation results take the probability weights into account and allow for correlation between error terms within schools.[5]

Table 1 reports descriptive statistics on observable teacher characteristics separately for math and science teachers. Mean differences are reported in the last column of table 1. The share of teachers with math or science majors naturally differs between those two groups. Apart from mean differences in majors only a few other variables are significantly different between the two groups.

To investigate which teachers teach using relatively more lecture style presentations and which students are more exposed to this method tables 2 and 3 display averages of student, school, class and teacher characteristics grouped by the share of lecture style presentation. The last column of tables 2 and 3 presents the mean differences between characteristics of students, schools, classes and teachers with more and less than the median share of time spent on lecture style presentation.

Table 2 reveals that the only student characteristic which shows significant differences between the two groups is the number of books that families have at home. Students who are taught with relatively less lecture style seem to have a little less books at home. Furthermore, the share of students in schools for which the principal reports very low involvement of parents in school activities is higher among students taught with relatively less lecture style presentations. Students exposed to below median lecture style teaching are also slightly more likely to be in tracked classes.

Table 3 reports teacher characteristics by the intensity of lecture style teaching. Students taught with relatively less lecture style teaching have a higher share of female teachers, more teachers who are at least 50, teachers who have the maximum number of years of teacher training, and teachers who have taken pedagogical or content knowledge classes in the last two years. All other teacher characteristics do not differ significantly between the two groups.

Tables 2 and 3 indicate some potential for selection into different teaching practices. To control for the confounding impact of these differences in observable characteristics, we include all variables presented in the tables above and additional information on teaching limits (see A-1 in the appendix) in our empirical analysis.

[5] 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.

4 Estimation Strategy

To estimate the effect of giving lecture style presentations relative to teaching based on problem solving we estimate a standard education production function:

$$Y_{ijk} = c_j + B'_{ijk}\beta_{1j} + S'_{ijk}\beta_{2j} + T'_{ijk}\beta_{3j} + Lecture'_{ijk}\beta_{4j} + \epsilon_{ijk}. \tag{1}$$

The test score, Y_{ijk}, of student i in subject j in school k is explained by student background characteristics, B_{ijk}, school characteristics, S_{ijk}, teacher and class characteristics, T_{ijk}, and the variable $Lecture_{ijk}$. The last variable constitutes the focus of this analysis. It represents teaching time spent on lecture style presentation relative to problem solving. The error term, ϵ_{ijk}, contains all unobservable influences on student test scores. In particular, it contains the effects of unobservable student, μ_i , teacher, ξ_j, and school characteristics, ν_k:

$$\epsilon_{ijk} = \mu_i + \xi_j + \nu_k + \psi_{ijk} \tag{2}$$

Estimating equation (1) by ordinary least squares produces biased estimates if unobserved school characteristics, ν_k, and $Lecture_{ijk}$ are correlated. This can be the case if the choice of the teaching practice is partly determined by the school and if there exists sorting of high ability students or effective teachers into schools.

To eliminate the effects of between-school sorting, we use school fixed effects, s_k, to exclude any systematic between-school variation in performance or teaching practice, whatever its source:

$$Y_{ijk} = c_j + B'_{ik}\beta_{1j} + s_k + T'_{ijk}\beta_{3j} + Lecture'_{ijk}\beta_{4j} + \mu_i + \xi_j + \psi_{ijk}. \tag{3}$$

The estimates produced by equation (3) could still be biased by within-school sorting wherever schools have more than one class per subject per grade. We therefore eliminate the influence on constant student traits by differencing between subjects:

$$\begin{aligned}\Delta Y_i \;&=\; c_m - c_s + B'_i(\beta_{1m} - \beta_{1s}) + S'_i(\beta_{2m} - \beta_{2s}) \\ &\quad + T'_{im}\beta_{3m} - T'_{is}\beta_{3s} + Lecture'_{im}\beta_{4m} - Lecture'_{is}\beta_{4s} + \eta_i\end{aligned} \tag{4}$$

where $\Delta Y_i = Y_{im} - Y_{is}$ and $\eta_i = \xi_m - \xi_s + \psi_{im} - \psi_{is}$.

In our headline specification we follow Dee (2005, 2007) by assuming that coefficients for each variable are equal across the two subjects:[6]

$$\Delta Y_i \;\; = \Delta T_i' \beta_3 + \Delta Lecture_i' \beta_4 + \eta_i. \tag{5}$$

The estimate of the effect of teaching practice on student achievement produced by equation (5) is not biased due to between or within school sorting of students based on unobservable student traits. We do, however, have to make the identifying assumption that unobservable teacher characteristics that directly influence student achievement are not related to the choice of the teaching method. In other words, η_i is uncorrelated with all other right-hand side variables. This is a strong assumption and we therefore refrain from interpreting β_4 as causal effect. We rather interpret β_4 as a measure for the link between a teaching practice and student achievement that is not driven by between or within school sorting of students. It might, however, be partly driven by sorting of teachers into a special teaching method based on unobservable teacher traits.

We evaluate the concern of selection on unobservables by borrowing a procedure from Altonji et al. (2005) which allows to evaluate the bias of the estimate under the assumption that selection on unobservables occurs to the same degree as selection on observables. As developed in the appendix the asymptotic bias of $\widehat{\beta_4}$ in our application is

$$Bias(\widehat{\beta_4}) = \frac{Cov(\Delta \widetilde{Lecture}, \eta)}{Var(\Delta \widetilde{Lecture})} = \frac{Cov(\Delta Lecture, \eta)}{Var(\Delta \widetilde{Lecture})} \tag{6}$$

where indices are omitted for simplicity and where $\Delta \widetilde{Lecture}$ is the residual of a linear projection of $\Delta Lecture$ on all other between subject differences of control variables, represented by ΔT. The second equality holds if the other controls (T) are orthogonal to η. The condition that selection on unobservables is equal to selection on observables can be stated as

$$\frac{Cov(\Delta T' \beta_3, \Delta Lecture)}{Var(\Delta T' \beta_3)} = \frac{Cov(\Delta Lecture, \eta)}{Var(\eta)} \tag{7}$$

Equation (7) can be used to estimate the numerator of the bias of $\widehat{\beta_4}$, once we have consistent estimates for β_3. Under the assumptions that the true effect of lecture style

[6] We do, however, estimate equation (4) as a robustness check.

teaching is zero and again that T is orthogonal to η, β_3 can be consistently estimated (see appendix).

The estimated bias displays the effect we would estimate even if the true effect was zero when selection on unobservables is as strong as selection on observables. In addition, we report the ratio of the estimated β_4 from equation (5) and the estimated bias giving a hint of how large selection on unobservables would have to be compared to selection on observables to explain the entire estimated effect. A value higher than one indicates that selection on unobservables needs to be stronger than selection on observables to explain the entire estimate, in case of a ratio lower than one already weaker selection on unobservables than on observables suffices to explain the entire estimate.

5 Results

Estimates of the effect of teaching practices based on the different methods advanced in Section 4 are presented in Tables 4 and 5. Each regression is performed at the level of the individual student and each of the estimations also takes into account the complex data structure produced by the survey design and the multi-level nature of the explanatory variables.

Table 4 reports results from estimating equation (1) and equation (3). We estimate both equations separately for math and science. Columns 1 and 3 present regressions results for math and science based on equation (1). These regressions include a complete set of student- and family-background variables, controls for teacher and class characteristics as well as characteristics of the school. Given the purpose of this study, only estimated coefficients for the teaching practice variable of interest and selected teacher characteristics are reported.

Our key variable of interest, teaching time devoted to lecture style presentation relative to time spent on problem solving, is estimated to have a positive impact on test scores in both subjects. In math the estimate is highly significant, while the estimate in science falls short of achieving statistical significance at any common significance level.

As discussed in the previous section, these results might be confounded by between school sorting based on unobservable characteristics of students. Column 2 and 4 therefore report estimation results based on equation (3), which includes school fixed effects. Lecture style presentation is now highly significant in science and the point estimate significantly increased compared to column 3. The estimate in math, however, did not change but lost its statistical significance due to increased standard errors.

To gain statistical power we pool both estimation samples and estimate equations (1) and (3) with the joint sample. This approach assumes that the effects of all right-hand side variables are identical in both subjects. Based on this estimation sample the relationship between more lecture style presentation and test scores is positive and significantly different from zero in both specifications.

The evidence presented in table 4 suggests a positive relation between more lecture style presentation and student achievement. However, within school selection of students based on unobservable student characteristics might drive this relationship. For instance, it is reasonable to assume that teachers adjust the use of teaching practices according to class composition and average student ability. We therefore difference out unobserved constant student traits by taking between-subject differences of test scores and all right-hand side variables as presented in equation (5).

Table 5 presents estimation results of the between-subject differences approach. We start out with a very basic specification without further controls in column 1 and successively add more controls that vary between subjects to account for subject-specific differences. In particular column 6 presents our headline results. In this specification we additionally control for all observable teacher and class characteristics presented in tables 1 and A-1. It is quite astonishing to see that adding more control variables in the between-subject specification leaves our estimate for the effect of more lecture style teaching almost unchanged. Moreover, in contrast to most other control variables the share of lecture style teaching is estimated to be statistically significant throughout all specifications presented in table 5. Apart from lecture style teaching, additional minutes per week spent teaching the subject to the class also seem to have a substantial positive effect on student test scores. This is a very intuitive result as it simply suggests that students learning is an increasing function of total teaching time.

The specifications in columns 2 to 6 also include other class activities, i.e. the share of total time in class spent on other activities apart from lecture style presentation or problem solving, as a control variable. Other class activities include time spent on any of the other five categories in question 20. In particular, these other activities include reteaching material and accountability measures like reviewing homework and test and quizzes. Naturally, not all students understand material when it is taught for the first time so that reviewing can be beneficial for student achievement. Moreover, previous research has shown that accountability measures might have positive effects for student achievement (Wenglinsky, 2000, 2002). The overall average effect of all other class activities is, however, estimated to be small and insignificant.

The estimate of teaching time devoted to lecture style presentation relative to time spent on problem solving decreases in comparison to the regression results presented in table 4. This indicates that within school sorting matters for the estimation of teachers' choice variables such as the degree of lecture style teaching. Estimates for the effect of more lecture style teaching on student learning in table 5 range from 0.14 to 0.1. Our headline estimate is reported in column 6 with an estimated size of 0.1.[7] This parameter suggests that shifting 10 p.p. of time from problem solving to teaching by giving lecture style presentations while holding the overall time devoted to these two activities constant is associated with an increase in student test scores of one percent of a standard deviation.

In turn, our results imply a negative correlation between more in-class problem solving and student achievement. This is consistent with the finding in Brewer and Goldhaber (1997) that more in-class problem solving for tenth grade students in math is related to lower test scores on a standardized test. Furthermore, the other commonly investigated teacher characteristics (e.g. gender, experience, credentials etc.) do not show significant or robust impacts on student achievement as can be seen in table 5. This is in line with previous findings in this literature and emphasizes the importance of the statistical significant relationship between more lecture style teaching and student achievement.

As pointed out in the previous section, estimates might still be biased due to selection of teachers into more (or less) lecture style teaching based on unobservable teacher characteristics. This concern is fostered by previous findings in the literature that emphasize the importance of unobservable teacher traits for student achievement. This raises the question: How can these results be interpreted?

The bias and ratio at the end of table 5 allow us to shed some light on the question of the influence of unobservables. The underlying assumption for the estimation of each bias is that selection on unobservables occurs to the same degree as selection on observables. In all columns the estimated bias is larger than the point estimate of the impact of lecture style teaching on student test scores. This is reflected in the ratios at the end of each column that are always smaller than one, indicating that selection on unobservables that is weaker than selection on observables suffices to explain the entire estimated coefficient. In our headline specification in column 6 selection on unobservables that is only 0.07 times

[7]It should be noted that the lecture style teaching variable is based on self-reported time-use by teachers, which is likely measured with error. If time-use is measured with classical measurement error, the estimate of the effect of lecture style teaching is biased towards zero. Hence, 0.1 might be interpreted as a lower bound for the true effect of lecture style teaching.

as strong as selection on observables would explain the entire estimated coefficient given that the true effect is zero. On the one hand, we have included a great amount of control variables so that we believe that selection on unobservables is likely weaker than selection on observables. On the other hand, only very little selection on unobservables compared to the selection on observables suffices to explain the entire effect. Given this uncertainty, we refrain from interpreting the results as evidence for a causal effect as the positive coefficient could also reflect selection of teachers with desirable unobserved characteristics into lecture style teaching.

This raises another question: Why would teachers with different desirable unobserved characteristics select different degrees of lecture style teaching compared to problem solving? While a reduced form approach of educational production cannot mirror the full complexity of the choices involved in the teaching process, we are, nevertheless, able to pin down the relationship between potential selection based on unobserved teacher traits and the causal effect of lecture style teaching as our estimation approach eliminated all other likely biases. If no selection based on unobservable teacher traits exists, our estimates speak for a positive effect of lecture style teaching. Our estimates might, however, be biased upwards if teachers with desirable unobserved characteristics more frequently base their instruction on lectures. Theoretically, this selection bias could be large enough to hide a true negative effect of lecture style teaching, which would imply that teachers with desirable unobserved characteristics predominately select themselves into an inferior teaching practice. This scenario, however, lacks any intuitional or theoretical support. We thus argue that this scenario is highly implausible and can be excluded, which allows a conservative interpretation of our results: We find no evidence for any detrimental effect of lecture style teaching on overall student learning.

It is important to stress that our results are limited to student achievement as measured by TIMSS test scores.[8] Moreover, our results for lecture style teaching are based on a comparison with teaching based on in-class problem solving. Loosely speaking our results suggest that on average lecture style teaching is at least not worse than teaching based on in-class problem solving. The comparison of average effects might, however, hide significant effect heterogeneities. Depending on the teacher, the students, the content taught, or other factors the one or the other teaching practice might be more effective. Furthermore, our information on teaching practices, which is based on in-class time use reported by teachers,

[8]To the extent that teaching methods cultivate general test-taking abilities differently or generate different long-term effects, a focus on other learning outcomes or on long-term effects might produce different results.

does not allow us to distinguish between different implementations of the teaching practices. One worry might be that especially the implementation of problem-based teaching differs substantially within our sample. Thus, while a certain teaching practice may be very effective if implemented in the correct way, an empirical analysis that is based on the actual average implementation of this teaching practice might not reveal any positive effects. Our results, therefore, do not call for more lecture style teaching in general. The results rather imply that the high share of traditional lecture style teaching in US middle schools is presumably less of a problem than often believed. The findings also suggest that simply reducing the amount of lecture style teaching and substituting it with more in-class problem solving without concern for how this is implemented is unlikely to raise overall student achievement in math and science.

6 Robustness Checks

This section tests the sensitivity of the results presented in section 5 with respect to other definitions of the lecture style variable and with respect to specifications allowing for heterogeneous effects. The results of these robustness checks are presented in tables 6 and 7.

As our grouping of the response categories available to the teacher in question 20 of the 2003 teacher questionnaires in TIMSS could be criticized, we provide evidence on the effect of interest based on different approaches to construct the lecture style variable. We test four alternative definitions of the lecture style variable with corresponding estimation results presented in each of the four columns of the upper panel of table 6.

In column 1 time spent re-teaching and clarifying content/procedures is included in the proxy for lecture style teaching. In column 2 taking tests or quizzes is added to the problem solving category. Hence, the lecture style teaching variable in column 2 is defined in relation to the enlarged definition of teaching based on problem solving. In column 3 we decompose the variable other class activities into its elements and separately control for each category. In column 4 lecture style is defined as the share of overall time in class spent on giving lecture style presentation.

The coefficients in the upper panel of table 6 reveal that redefining our key variable of interest does not change the estimated impact of more lecture style teaching on student achievement. While the main purpose of this study is to analyze the teaching of new material by giving lecture style presentations rather than by letting pupils solve problems, it is reassuring to see that increasing the total amount of time in class devoted to lecture style presentations (in contrast to all other in-class activity) is also associated with higher student achievement.

Additionally, we present evidence for various sub-samples in the middle panel of table 6. In column 1 and 2 we estimate equation (5) for pupils with the same peers in math and science and pupils with different peers, respectively. This distinction is motivated by the concern that the main effect might be driven by differences in the classroom composition. In the sub-sample with identical peers in both subjects our within-student between-subject identification strategy takes care of any potential peer effects. For students with the same peers in both subjects, shifting 10 p.p. of teaching time from problem solving to lecture style teaching is associated with an increase of almost 4 percent of a standard deviation. The estimate in the sub-sample with different peers decreases to 0.08 and lacks statistical significance. The results indicate that peer effects do not drive the positive coefficient of lecture style teaching.

Column 3 and 4 of the middle panel of table 6 report estimates separately for students in schools where either no or both subjects are tracked by ability and for students in schools where tracking on ability exists in only one of the two subjects. This distinction is motivated by the consideration that tracking on ability might induce teachers to choose different degrees of lecture style teaching. The results indicate that the positive association between more lecture style teaching and student achievement holds in both types of schools. The point estimate is the same for the two groups. In schools with differential tracking policies, however, it is not statistically significant.

In the lower panel of table 6 we investigate subject-specific effects. Column 1 and 2 present estimation results from estimating versions of equation (4), where we abandon the assumption that coefficients for each right-hand side variable are equal across subjects. As all science variables enter negatively on both sides of equation (4), a negative coefficient for any variable in science masks a positive relationship between the variable and the science test score. All estimates thus have the expected signs. They are not statistically significant for science, though. We thus find evidence for a stronger effect of lecture style teaching in math. A possible interpretation of the differential effects in the two subjects is that science with its natural emphasis on experimentation might just be better suited for problem solving.

So far all specifications measured time devoted to certain class activities in shares while additionally controlling for total teaching time. The specification in shares is mainly motivated by the data itself as TIMSS asks teachers to report what percentage of time in a typical week of the specific subject's lessons students spend on various activities. The data also contains information on the total time in minutes per week teachers teach math or science to the class. Thus, we can also construct a proxy for minutes per week devoted to each activity based on the information on total time per week and the shares reported by the teachers. While this constructed measure presumably suffers from larger measurement error than the reported shares alone, we nevertheless estimate a version of equation 5 that is based on variables indicating various teaching practices measured in minutes per week as a further robustness check.

The results of this robustness check can be seen in table 7. All specifications measure class activities in minutes per week. Otherwise the specifications are identical to the specification estimated in column 6 of table 5. In column 1 of table 7, minutes spent per week with lecture style teaching and minutes spent on other class activities are included in the specification while minutes spent on problem solving are excluded as a reference category. One additional minute per week spent on lecture style teaching instead of problem solving is associated with an increase in test scores of 0.0007 percent of a standard deviation. Column 2 of table 7 reveals that increasing the overall time devoted to lecture style teaching leads to an increase in test scores of the same magnitude as in column 1 when explicitly including the two problem solving categories and the other class activities. Columns 3 and 4 control again for total teaching time, but use either problem solving with the teachers guidance (column 3) or problem solving without guidance (column 4) as reference categories. The point estimate of lecture style teaching remains positive in both specifications, but loses significance. The point estimates for the two problem solving categories in columns 3 and 4 suggest that problem solving with guidance is slightly more effective than problem solving without guidance. The difference is, however, not statistically significant.

In sum, we find positive relationships between more lecture style teaching and student achievement in all robustness analyses. The magnitude of the estimated effects varies between specifications and between sub-samples, with insignificant point estimates in some specification. Importantly, we do not find evidence for any detrimental effect of lecture style teaching in any specification.

7 Conclusion

Existing research on teacher quality allows two conclusions: First, there exists a large variation in teachers' ability to improve student achievement. Second, this variation cannot be explained by common, observable teacher characteristics. The results presented in this study confirm that these observable teacher characteristics have little potential for explaining the variation in student achievement. We provide, however, new evidence on a significant link between teaching practice and student achievement.

The specific teaching practice variable analyzed in this paper is the share of teaching time devoted to lecture style presentation (in contrast to in-class problem solving). We construct this variable based on information on in-class time use provided by teachers in the 2003 wave of the Trends in International Mathematics and Science Study (TIMSS) in US schools. Exploiting between-subject variation to control for unobserved student traits and estimating a reduced form educational production function, we find that a 10 percentage point shift from problem solving to lecture style presentation results in an increase in student achievement of about one percent of a standard deviation. We further show that this result is extremely robust to definitional changes in the construction of the main variable of interest as well as to specifications allowing for heterogeneous effects.

This finding suggests that students taught by teachers, who devote more teaching time to lecture style presentation rather than letting students solve problems on their own or with the teacher's guidance, learn more (in terms of competencies tested in the TIMSS student achievement test). This result stands in contrast to constructivist theories of learning. It is, however, in line with previous findings in the literature (Brewer and Goldhaber, 1997) showing that instruction in small groups and emphasis on problem solving lead to lower student test scores.

We emphasize, however, that our results demand a careful interpretation and need to be taken for what they are: Evidence for a positive association between more time devoted to lecture style teaching and student achievement that is neither driven by selection of particular students into schools or classes nor by selection of teachers based on various observable characteristics into a particular teaching method. However, selection based on unobservable teacher characteristics remains a worry. Following the method developed in Altonji et al. (2005), we show that only a relatively small selection based on unobservables suffices to explain the entire estimated coefficient. We thus refrain from formulating any policy conclusions that call for more lecture style teaching in general.

We are nevertheless able to draw an important conclusion about the nature of the causal effect of lecture style teaching on student achievement as we eliminated any potential biases arising from sorting of students, differences in schools and observable differences in teacher traits in our empirical approach. The existence of a sizeable negative causal effect of lecture style teaching would only be consistent with our results if teachers with favorable unobserved characteristics predominantly select themselves into an inferior teaching practice. Such a scenario, however, lacks any intuitional and theoretical support. We can thus exclude the possibility of a sizeable detrimental effect of lecture style teaching in math and science instruction on overall student achievement in US middle schools.

We believe that this result is relevant for the debate on optimizing the teaching process. Various dimensions of teaching practices have been shown to matter for student achievement. Moreover, the low-cost implementation of changes in the teaching process compared to other policy measures designed to foster student learning makes improvements in the teaching process particularly appealing. There exists, however, little consensus about what measures could improve the teaching process. Reducing the amount of traditional instruction based on lecture style teaching is typically a key candidate. Lectures are potentially connected with many disadvantages and might therefore be an inferior teaching method. National standards (NCTM, 1991; National Research Council, 1996) also advocate engaging students more in hands-on learning activities and group work but traditional lecture and textbook methodologies remain dominant in science and mathematics instruction in US middle schools. This raises the concern that the high share of total teaching time devoted to traditional lecture presentations has a detrimental effect on overall student learning in US middle schools. Our results, however, show that there exists no empirical support for this concern. Moreover, while newer teaching methods might be beneficial for student achievement if implemented in an ideal way, our findings imply that simply inducing teachers to shift time in class from lecture style presentations to problem solving without concern for how this is implemented contains little potential to increase student achievement. On the contrary, our results indicate that there might even be adverse effects on student learning.

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.

Altonji, J., Elder, T., and Taber, C. (2005). Selection on observed and unobserved variables: assessing the effectiveness of catholic schools. *Journal of Political Economy*, 113(1):151–184.

Aslam, M. and Kingdon, G. (2007). What can Teachers do to Raise Pupil Achievement? The centre for the study of African economies working paper series.

Banerjee, A. V., Cole, S., Duflo, E., and Linden, L. (2007). Remedying education: Evidence from two randomized experiments in India. *The Quarterly Journal of Economics*, 122(3):1235–1264.

Barrow, L., Markman, L., and Rouse, C. E. (2009). Technology's edge: The educational benefits of computer-aided instruction. *American Economic Journal: Economic Policy*, 1(1):52–74.

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

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.

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.

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

Gordon, R., Kane, T. J., and Staiger, D. O. (2006). The hamilton project: Identifying effective teachers using performance on the job. Report Washington, DC, The Brookings Institution.

Kane, T. J., Rockoff, J. E., and Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York city. *Economics of Education Review*, 27(6):615–631.

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

Machin, S. and McNally, S. (2008). The literacy hour. *Journal of Public Economics*, 92(5-6):1441–1462.

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.

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.

National Research Council (1996). National science education standards. Report, Washington, DC.

NCTM (1991). Professional standards for teaching mathematics. Report, National Council of Teachers of Mathematics, Reston, VA.

Newmann, F. M., Bryk, A. S., and Nagaoka, J. (2001). Authentic intellectual work and standardized tests: Conflict or coexistence? 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.

Rouse, C. E. and Krueger, A. B. (2004). Putting computerized instruction to the test: a randomized evaluation of a “scientifically based” reading program. *Economics of Education Review*, 23(4):323–338.

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

Seidel, T. and Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77(4):454.

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.

Weiss, I. R. (1997). The status of science and mathematics teaching in the United States: Comparing teacher views and classroom practice to national standards. NISE Brief 3, National Institute for Science Education, Madison: University of Wisconsin-Madison.

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

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- Teacher variables

	Math		Science		
	303 teachers		355 teachers		
Variable	Mean	SD	Mean	SD	Difference
In-class time use					
Lecture style teaching	0.320	0.187	0.374	0.202	-0.054***
Other class activities	0.428	0.119	0.446	0.161	-0.018
Total teaching time ($\frac{min}{week}$)	226.27	43.89	223.72	47.34	2.55
Teacher variables					
Teacher is female	0.649	0.473	0.540	0.496	0.109**
Full teaching certificate	0.970	0.163	0.957	0.188	0.013
Major in math	0.473	0.492	0.099	0.294	0.374***
Major in science	0.146	0.348	0.584	0.486	-0.438***
Major in education	0.598	0.483	0.456	0.491	0.142***
Teacher younger than 30	0.119	0.322	0.143	0.349	-0.024
Teacher aged 30-39	0.273	0.443	0.223	0.414	0.050
Teacher aged 40-49	0.293	0.452	0.335	0.469	-0.042
Teaching experience $<$ 1 year	0.043	0.201	0.042	0.199	0.001
Teaching experience 1-5 years	0.178	0.370	0.224	0.404	-0.046
Teacher training 0 years	0.102	0.301	0.154	0.359	-0.051**
Teacher training 1 year	0.578	0.491	0.523	0.497	0.055
Teacher training 2 years	0.209	0.404	0.193	0.392	0.016
Teacher training 3 years	0.039	0.192	0.048	0.213	-0.009
Teacher training 4 years	0.056	0.228	0.035	0.184	0.021
Teacher training 5 years	0.008	0.090	0.039	0.193	-0.031**
Motivation					
Pedagogy classes in last 2 years	0.748	0.431	0.648	0.472	0.100***
Subject content classes in last 2 years	0.840	0.364	0.827	0.374	0.014
Subject curriculum classes in last 2 years	0.830	0.372	0.853	0.349	-0.023
Subject related IT classes in last 2 years	0.729	0.441	0.803	0.393	-0.074**
Subject assessment classes in last 2 years	0.756	0.426	0.649	0.471	0.107***
Classes on improving student's skills last 2 years	0.759	0.424	0.766	0.418	-0.007
Working hours scheduled per week	21.12	8.28	20.16	7.29	0.960
Weekly hours spent on lesson planning	3.70	2.71	4.68	3.28	-0.976***
Weekly hours spent on grading	5.25	3.93	6.08	4.41	-0.830**
Teaching requirements					
Requirement probationary period	0.502	0.493	0.496	0.479	0.007
Requirement licensing exam	0.526	0.493	0.558	0.479	-0.032
Requirement finished Isced5a	0.891	0.307	0.824	0.371	0.067**
Requirement minimum education classes	0.833	0.368	0.777	0.399	0.056
Requirement minimum subject specific classes	0.799	0.395	0.744	0.420	0.056

Note: Probability weights and within school correlation are taken into account when estimating means and standard deviations. Teacher variables are weighted by the number of students taught by each teacher.

Table 2: Student, School and Class Variables by Intensity of Lecture Style Teaching

Variable	**Lecture Style <= median** Mean	SD	**Lecture Style > median** Mean	SD	Difference
Student					
Age	14.23	0.470	14.22	0.451	0.014
Born in first 6 months	0.495	0.499	0.501	0.499	-0.005
Girl	0.522	0.499	0.522	0.500	0.000
Number books at home: 11-25	0.173	0.376	0.175	0.378	-0.002
Number books at home: 26-100	0.278	0.445	0.282	0.448	-0.004
Number books at home: 101-200	0.173	0.375	0.185	0.387	-0.013*
Number books at home: >200	0.250	0.430	0.248	0.430	0.002
Parental Education lower secondary	0.058	0.211	0.057	0.208	0.001
Parental Education upper secondary	0.253	0.392	0.248	0.388	0.005
Parental Education post secondary voc/technical	0.088	0.254	0.094	0.264	-0.006
Parental Education university	0.571	0.446	0.577	0.445	-0.006
Number people at home 3	0.170	0.372	0.165	0.368	0.005
Number people at home 4	0.350	0.472	0.347	0.472	0.003
Number people at home 5	0.251	0.429	0.255	0.432	-0.004
Number people at home 6	0.104	0.303	0.108	0.307	-0.003
Number people at home 7	0.042	0.198	0.046	0.207	-0.004
Number people at home 8 or more	0.039	0.193	0.039	0.192	0.000
Never speaks English at home	0.048	0.213	0.044	0.205	0.004
Sometime speaks English at home	0.093	0.289	0.099	0.298	-0.006
Almost always speaks English at home	0.843	0.361	0.845	0.360	-0.002
Immigrant	0.144	0.349	0.141	0.347	0.003
School					
Community, 3001-15,000 people	0.228	0.403	0.188	0.375	0.041
Community, 15,001-50,000 people	0.293	0.435	0.330	0.454	-0.037
Community, 50,001 - 100,000 people	0.118	0.309	0.121	0.314	-0.003
Community, 100,001 - 500,000 people	0.130	0.321	0.137	0.331	-0.007
Community, more than 500,000 people	0.111	0.299	0.125	0.319	-0.014
Parental involvement in school - very low	0.204	0.384	0.130	0.319	0.074**
Parental involvement in school - low	0.324	0.442	0.389	0.469	-0.065
Parental involvement in school - medium	0.316	0.441	0.300	0.440	0.016
Parental involvement in school - high	0.101	0.285	0.118	0.311	-0.017
Class variables					
Total teaching time ($\frac{min}{week}$)	227.45	43.03	221.93	48.60	5.52
Class size	24.03	7.20	23.95	6.90	0.073
Tracked according to ability	0.394	0.469	0.319	0.452	0.075*

Note: Lecture style measures the share of teaching time devoted to lecture style teaching instead of teaching based on problem solving. Median refers to the median of lecture style teaching. Probability weights and within school correlation are taken into account when estimating means and standard deviations.

Table 3: Teacher Variables by Intensity of Lecture Style Teaching

Variable	**Lecture Style <= median** Mean	SD	**Lecture Style > median** Mean	SD	Difference
Teacher variables					
Teacher is female	0.642	0.475	0.535	0.496	0.107**
Full teaching certificate	0.977	0.135	0.947	0.216	0.031
Major in math	0.303	0.455	0.265	0.432	0.039
Major in science	0.349	0.471	0.385	0.477	-0.036
Major in education	0.537	0.493	0.515	0.490	0.022
Teacher younger than 30	0.134	0.339	0.127	0.331	0.007
Teacher aged 40-49	0.272	0.442	0.367	0.480	-0.096**
Teacher at least 50	0.347	0.473	0.256	0.434	0.091**
Teaching experience < 1 year	0.041	0.197	0.045	0.205	-0.004
Teaching experience 1-5 years	0.180	0.372	0.229	0.406	-0.049
Teacher training 0 years	0.127	0.331	0.130	0.334	-0.003
Teacher training 1 year	0.537	0.496	0.568	0.492	-0.031
Teacher training 2 years	0.191	0.391	0.213	0.407	-0.022
Teacher training 3 years	0.050	0.216	0.036	0.185	0.013
Teacher training 4 years	0.050	0.216	0.040	0.195	0.009
Teacher training 5 years	0.036	0.186	0.008	0.089	0.028**
Teacher motivation					
Pedagogy classes in last 2 years	0.749	0.429	0.633	0.476	0.116***
Subject content classes in last 2 years	0.865	0.338	0.792	0.402	0.073**
Subject curriculum classes in last 2 years	0.860	0.342	0.818	0.381	0.042
Subject related IT classes in last 2 years	0.754	0.427	0.780	0.409	-0.025
Subject assessment classes in last 2 years	0.728	0.441	0.671	0.464	0.057
Classes on improving student's skills last 2 years	0.768	0.418	0.757	0.424	0.011
Working hours scheduled per week	20.50	8.09	20.82	7.45	-0.322
Weekly hours spent on lesson planning	4.12	2.96	4.29	3.14	-0.170
Weekly hours spent on grading	5.67	4.45	5.66	3.85	0.013
Teaching requirements					
Requirement probationary period	0.496	0.487	0.503	0.484	-0.008
Requirement licensing exam	0.556	0.487	0.523	0.486	0.033
Requirement finished ISCED 5a	0.857	0.343	0.858	0.340	-0.001
Requirement minimum education classes	0.826	0.369	0.779	0.402	0.047
Requirement minimum subject specific classes	0.797	0.392	0.740	0.425	0.056

Note: Lecture style measures the share of teaching time devoted to lecture style teaching instead of teaching based on problem solving. Median refers to the median of lecture style teaching. Probability weights and within school correlation are taken into account when estimating means and standard deviations. Teacher variables are weighted by the number of students taught by each teacher.

Table 4: Estimation Results OLS

	Math1	Math2	Science1	Science2	Pooled1	Pooled2
Lecture style teaching	.514**	.431	.207	.690**	.380***	.293***
	(.20)	(.40)	(.13)	(.29)	(.11)	(.09)
Other class activities	–.0703	–.701	.0416	–.461	.0215	–.322**
	(.30)	(.52)	(.16)	(.30)	(.14)	(.14)
Total teaching time ($\frac{min}{week} * 10^{-3}$)	.00364	.328	–.329	–.157	–.0639	.0702
	(.66)	(2.98)	(.56)	(.87)	(.41)	(.45)
Female teacher	–.156**	–.171	–.0747	–.0224	–.0982**	–.0187
	(.06)	(.12)	(.06)	(.10)	(.05)	(.04)
Teacher younger than 30	.125	.0740	.116	.256*	.110	.0544
	(.14)	(.24)	(.09)	(.14)	(.08)	(.08)
Teacher's age 40-49	.0545	.0753	.0545	.127	.0509	.00123
	(.08)	(.18)	(.07)	(.10)	(.05)	(.05)
Teacher older than 50	.0242	.113	.0833	–.0231	.0337	.0158
	(.08)	(.18)	(.07)	(.13)	(.06)	(.06)
Teaching experience <1 year	–.513***	–.878***	–.116	–.0216	–.266**	–.160*
	(.15)	(.24)	(.17)	(.18)	(.11)	(.09)
Teaching experience 1-5 years	–.202**	–.303*	–.0160	–.0320	–.103	–.0773
	(.10)	(.18)	(.07)	(.10)	(.07)	(.07)
Teaching certificate	–.368**	–1.229***	.0572	–.358	–.0395	–.132
	(.17)	(.32)	(.18)	(.38)	(.14)	(.16)
Constant	.898	1.891	–.0151	2.340***	.345	.345
	(.55)	(1.30)	(.56)	(.73)	(.49)	(.42)
Student Background	Yes	Yes	Yes	Yes	Yes	Yes
School variables	Yes	No	Yes	No	Yes	No
School fixed effects	No	Yes	No	Yes	No	Yes
Teacher variables	Yes	Yes	Yes	Yes	Yes	Yes
Class variables	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6310	6310	6310	6310	12620	12620
R^2	.313	.518	.337	.501	.309	.476

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

Note: Dependent variable is the standardized student test score. Lecture style measures the share of teaching time devoted to lecture style teaching instead of teaching based on problem solving. Student background includes students' gender, age in years, a dummy if born in first 6 months of the year, number of books at home, English spoken at home, migration background, household size and parental education. School variables include dummies capturing different levels of parental involvement in school activities and dummies for community size. Not reported teacher variables are teacher's major in math, science and education and years of teacher training. Class variables are class size, and a tracking indicator. Imputation indicators are included in all regressions. Standard errors clustered at the school level in parentheses.

Table 5: Estimation Results First Difference

	1	2	3	4	5	6
Lecture style teaching	.112*	.118**	.139***	.132***	.126**	.127**
	(.06)	(.06)	(.05)	(.05)	(.05)	(.05)
Other class activities		−.0115	.0195	.0290	.0360	.0304
		(.07)	(.06)	(.06)	(.06)	(.06)
Total teaching time ($\frac{min}{week} * 10^{-3}$)		.438	.498*	.380*	.402*	.524**
		(.28)	(.26)	(.22)	(.23)	(.23)
Teacher female			−.000597	−.00723	−.0107	−.0117
			(.02)	(.02)	(.02)	(.02)
Teaching certificate			.0296	.0378	.0747	.0740
			(.06)	(.11)	(.12)	(.12)
Teacher younger than 30			−.0413	−.0583*	−.0490	−.0433
			(.04)	(.03)	(.04)	(.04)
Teacher's age 40-49			.0179	.0472*	.0606**	.0516*
			(.03)	(.03)	(.03)	(.03)
Teacher at least 50			.0248	.0456*	.0496*	.0467
			(.03)	(.03)	(.03)	(.03)
Teaching experience <1 year			−.00302	−.00550	−.00513	−.0202
			(.04)	(.05)	(.05)	(.05)
Teaching experience 1-5 years			.0266	.0355	.0375	.0272
			(.03)	(.03)	(.03)	(.03)
Constant	−.0103	−.0253	.00360	.000995	−.00663	−.0210
	(.01)	(.02)	(.03)	(.02)	(.03)	(.02)
Class variables	No	No	Yes	Yes	Yes	Yes
Teacher variables	No	No	Yes	Yes	Yes	Yes
Limit to teach	No	No	No	Yes	Yes	Yes
Motivation	No	No	No	No	Yes	Yes
Teaching requirements	No	No	No	No	No	Yes
Observations	6310	6310	6310	6310	6310	6310
R^2	.002	.005	.014	.029	.031	.035
Bias[a)]				1.265	1.578	1.723
Ratio [b)]				.104	.080	.0738

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

Note: a) Bias estimated as in equation (6) using condition (7) b) Ratio of the coefficient of lecture style presentation and the bias. Dependent variable is the within student difference of standardized math and science test scores. All teacher variables are included as within student between subject differences. Lecture style measures the share of teaching time devoted to lecture style teaching instead of teaching based on problem solving. Variables included in Motivation, Teacher variables and Teaching requirements are shown in table 1. Variables included in Class variables and teaching limits are shown in table A-1. Imputation indicators included in all but the first two columns. Standard errors clustered at the school level in parentheses.

Table 6: Robustness Checks I

Other Definitions	Def 1	Def 2	Def 3	Def 4
Lecture style teaching	.145**	.161**	.134**	.156*
	(.06)	(.06)	(.05)	(.08)
Observations	6310	6310	6310	6310
R^2	.035	.036	.036	.035

Subsamples	Same Peers	Diff Peers	No Track	Track
Lecture style teaching	.355***	.0821	.116*	.116
	(.12)	(.06)	(.07)	(.09)
Observations	2205	4105	3529	2292
R^2	.096	.046	.065	.071

Heterogenous effects	Diff	Background
Lecture style teaching	.203***	.130*
(math)	(.07)	(.07)
Lecture style teaching	−.105	−.111
(science)	(.07)	(.07)
Observations	6310	6310
R^2	.065	.081

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

Note: Dependent variables in all panels and columns are the within student between subject differences in standardized test scores. All teacher variables, class variables, motivation and teaching limits are included as controls. Upper panel: In Def 1 time spent re-teaching and clarifying content/procedures is added to lecture style teaching. In Def 2 taking tests or quizzes is added to problem solving. Thus, lecture style teaching in column 2 is defined in relation to the enlarged definition of teaching based on problem solving. In Def 3 the variable 'other class activities' is decomposed into its elements and these are separately included to control for each category. Def 4 takes time spent on giving lecture style presentation in relation to all other time-use categories (problem solving + other class activities). Middle panel: Separate estimation for different sub-samples: Column 1 only students with same classmates in both subjects, column 2 students with different classmates. Column 3 students who are tracked according to ability in either both or none of the two subjects, column 4 students who are tracked in at least one of the two subjects. Lower panel: Column 1 and 2 allow different coefficients in the two subjects. A negative coefficient for science variables stands for a positive association with the dependent variable. Column 2 additionally includes student background as controls. Imputation indicators are included in all estimations. Standard errors clustered at the school level in parentheses.

Table 7: Robustness Checks II: Absolute Time Specification

Variables are measured as minutes per week spent teaching with specific teaching practice.

	(1)	(2)	(3)	(4)
Lecture style teaching	.733*	.751*	.653	.826
	(.43)	(.41)	(.48)	(.53)
Problem solving with guidance		.184		.174
		(.37)		(.54)
Problem solving w\o guidance		.0577	−.174	
		(.42)	(.54)	
Other class activities	.444	.551*	.362	.536
	(.32)	(.30)	(.36)	(.46)
Total teaching time	.150		.236	.0624
	(.29)		(.37)	(.42)

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

Note: All teaching practice variables are rescaled as minutes per week divided by (10^3). Column (1) groups time spent on all activities but lecture style teaching and problem solving, problem solving with and without guidance as reference category. Column (2) includes all in-class time use categories, excluded is total minutes per week. Column (3) uses only problem solving with guidance as reference category. Column (4) uses only problem solving without guidance as reference category. All specifications are otherwise identical to the specification estimated in Column (6) of table 5. Standard errors clustered at the school level in parentheses.

Appendix

Selection on unobservables following Altonji et al. (2005)

Formally, in our application the assumption that selection on unobservables occurs to the same degree as selection on observables as imposed by Altonji, Elder and Taber (2005) can be stated as:

$$
\begin{aligned}
Proj(\Delta Lecture|\Delta T\beta_3, \eta) \quad &= \quad \phi_0 + \phi_{\Delta T'\beta_3}\Delta T'\beta_3 + \phi_\eta \eta && (8) \\
\text{with} \qquad \phi_{\Delta T'\beta_3} \quad &= \quad \phi_\eta && (9)
\end{aligned}
$$

Where $\Delta Lecture$ captures the between subject differences in the teaching time devoted to lecture style presentation relative to problem solving and β_4 indicates its coefficient while T includes all other k control variables (teacher characteristics and effective teaching time as well as class characteristics) and β_3 is a $kx1$ vector of coefficients. When $\Delta T'\beta_3$ is orthogonal to η the assumption (9) is equal to

$$
\frac{Cov(\Delta T'\beta_3, \Delta Lecture)}{Var(\Delta T'\beta_3)} = \frac{Cov(\eta, \Delta Lecture)}{Var(\eta)} \tag{10}
$$

We proceed now to answer the question how large selection on unobservables relative to selection on observabels would have to be in order to explain the entire estimate of β_4 under the assumption that the true β_4 is 0. Following Altonji et al. (2005) we regress $\Delta Lecture$ on ΔT, to get

$$
\Delta Lecture = \Delta T'\delta + \Delta\widetilde{Lecture}.
$$

Plugging this into our equation (5) yields:

$$
\Delta Y \quad = c + \Delta T'(\beta_3 + \delta * \beta 4) + \Delta\widetilde{Lecture}^{'}\beta_4 + \eta \tag{11}
$$

As $\Delta\widetilde{Lecture}$ is by construction orthogonal to ΔT the probability limit of $\widehat{\beta}_4$ can be written as

$$
plim \ \ \widehat{\beta}_4 = \beta_4 + \frac{Cov(\Delta\widetilde{Lecture}, \eta_i)}{Var(\Delta\widetilde{Lecture})}
$$

where

$$
\frac{Cov(\Delta\widetilde{Lecture}, \eta)}{Var(\Delta\widetilde{Lecture})} = \frac{Cov(\Delta Lecture, \eta)}{Var(\Delta\widetilde{Lecture})}
$$

as ΔT is orthogonal to η.

To estimate the numerator of the bias we can use the equality (10):

$$\frac{Cov(\Delta T'\beta_3, \Delta Lecture)}{Var(\Delta T'\beta_3)} * Var(\eta).$$

For this however, we need a consistent estimate of β_3 which we obtain by estimating equation (11) under the assumption that $\beta_4 = 0$.

Table A-1: Descriptive Statistics- Class Characteristics

	Math		**Science**		
	359 classes		734 classes		
Variable	Mean	SD	Mean	SD	Difference
Class variables					
Class size	23.46	6.54	24.57	7.51	-1.11**
Tracked according to ability	0.550	0.480	0.171	0.363	0.379***
Teaching limits (reference not at all/not applicable)					
Differing academic ability - a little	0.339	0.469	0.340	0.473	-0.002
Differing academic ability - some	0.330	0.466	0.321	0.466	0.008
Differing academic ability - a lot	0.204	0.399	0.174	0.378	0.031
Wide range of backgrounds - a little	0.308	0.456	0.277	0.447	0.031
Wide range of backgrounds - some	0.205	0.399	0.238	0.425	-0.032
Wide range of backgrounds - a lot	0.060	0.234	0.078	0.267	-0.018
Special need students - a little	0.309	0.457	0.328	0.469	-0.019
Special need students - some	0.147	0.350	0.184	0.387	-0.037
Special need students - a lot	0.064	0.243	0.081	0.272	-0.016
Shortage computer hardware - a little	0.140	0.343	0.236	0.423	-0.097***
Shortage computer hardware - some	0.197	0.396	0.207	0.404	-0.011
Shortage computer hardware - a lot	0.110	0.310	0.189	0.390	-0.079***
Shortage computer software - a little	0.168	0.371	0.294	0.454	-0.125***
Shortage computer software - some	0.146	0.350	0.198	0.397	-0.052
Shortage computer software - a lot	0.145	0.349	0.174	0.378	-0.029
Shortage support pc use - a little	0.181	0.380	0.217	0.411	-0.036
Shortage support pc use - some	0.148	0.351	0.185	0.387	-0.037
Shortage support pc use - a lot	0.089	0.282	0.137	0.343	-0.048*
Shortage of textbooks - a little	0.055	0.225	0.088	0.283	-0.034
Shortage of textbooks - some	0.045	0.205	0.044	0.205	0.001
Shortage of textbooks - a lot	0.011	0.103	0.083	0.275	-0.072***

Table A-2: Descriptive Statistics- Class Characteristics (cont.)

	Math		**Science**		
	359 classes		734 classes		
Variable	Mean	SD	Mean	SD	Difference
Shortage instructional equipment - a little	0.180	0.380	0.314	0.463	-0.134***
Shortage instructional equipment - a some	0.123	0.326	0.193	0.394	-0.070**
Shortage instructional equipment - a lot	0.038	0.190	0.141	0.347	-0.103 ***
Shortage demonstrative equipment - a little	0.253	0.431	0.318	0.465	-0.065
Shortage demonstrative equipment - some	0.117	0.318	0.196	0.396	-0.080**
Shortage demonstrative equipment - a lot	0.044	0.203	0.189	0.391	-0.146***
Inadequate physical facilities - a little	0.148	0.352	0.219	0.413	-0.071*
Inadequate physical facilities - some	0.051	0.219	0.158	0.364	-0.107***
Inadequate physical facilities - a lot	0.030	0.169	0.131	0.337	-0.101***
High student teacher ratio - a little	0.230	0.417	0.292	0.454	-0.062
High student teacher ratio - some	0.132	0.335	0.204	0.402	-0.071**
High student teacher ratio - a lot	0.091	0.285	0.129	0.334	-0.038

Note: Probability weights and within school correlation are taken into account when estimating means and standard deviations. Class variables are weighted by the number of students in each class.