Lecturing Style Teaching and Student Performance

Chris Van Klaveren

TIER WORKING PAPER SERIES
TIER WP 10/16

Lecturing Style Teaching and Student Performance

Chris van Klaveren*

Abstract

Dutch teachers tend to give fewer lectures in front of the class, and instead 'choose' for a more personal approach, because it is believed that this positively affects student performance. However, the downside of a more personal approach is that it is time intensive and possibly eliminates the complementary and scale effects of giving lectures in front of the class.

This study examines whether the share of time that teachers spend on lecturing style teaching influences the cognitive performance of Dutch students. We find no relationship between lecturing style teaching and student performance. Hence, our results do not support the idea that lecturing style teaching is old fashioned or that a more personal teaching style would be beneficial for the cognitive performance of students.

JEL Codes: I21, C23

Keywords: Lecturing Styles, Teacher Quality, Student Performance

*University Maastricht, TIER, Faculty of Economics and Business Administration, P.O. BOX 616, 6200 MD Maastricht, The Netherlands. Email address of the corresponding author: cp.vanklaveren@maastrichtuniversity.nl. I would like to thank the people at TIER, and in particular Wim Groot, Henriette Maassen van den Brink, Martina Meelissen, Hessel Oosterbeek, Erik Plug and Dinand Webbink for their valuable remarks.

1 Introduction

Policy makers, parents and teachers agree that better teachers produce better students. However, from a scientific point of view, it is not so clear what defines a better teacher. It is believed, for example, that more experienced and higher educated teachers are better qualified to teach, but there is no unambiguous scientific evidence that these characteristics are causally related to student performance (see, among others, Hanushek, Kain, and Rivkin, 2005, Hanushek and Rivkin, 2006, Clotfelter, Ladd, and Vigdor, 2006, Clotfelter, Ladd, and Vigdor, 2007, Aaronson, Barrow, and Sander, 2007, Rothstein, 2008). The only strong and consistent empirical finding is that students perform less well when teachers have less than two years of teaching experience. However, even this finding has not so much to do with teacher quality, but has more to do with the not so productive 'freshman' period every teacher goes through.

In the Netherlands, teachers tend to give less lectures in front of the class, and instead 'choose' for a more personal approach, because it is believed that this affects the cognitive performance of students positively. The underlying thought is that the more personal approach allow teachers to better adjust their teaching method to the needs of the individual student, which on its turn increases student performance. However, the possible downside of a more personal approach is that it is time intensive and eliminates the complementary effects of lecturing style teaching. Therefore, a more individualistic teaching style may also be a less efficient teaching style.

This study examines whether the share of time that teachers spend on lecturing style teaching influences the cognitive performance of Dutch students. By focusing on the things that teachers do in class, rather than on the characteristics they possess, we obtain insights on whether lecturing style teaching is old fashioned and if a more personal teaching style would be beneficial for the cognitive performance of students.

For this purpose we exploit the TIMSS 2003 data for Dutch students who are in their second year of secondary education. These data contain information on the math and physics performance of students as well as information on student, teacher, school and class characteristics. Furthermore there is detailed information on how teachers fill in the students' day with respect to their math and physics lectures. Using this information we construct a lecturing style variable that represents the share of time that teachers spent on lecturing in front of the class.

There are, as far as we know, two studies that examine how the time that teachers allocate to different activities influences student performance. Aslam and Kingdon (2007) examine how student performance is influenced by several teacher activities and find that lesson planning, involving students by asking questions during class and quizzing them on past material, all substantially benefit pupil learning. Schwerdt and Wuppermann (2009) examine how the time that teachers spend on lecturing style teaching influences the performance of U.S. students and find that students of teachers who lecture more in front of the class perform slightly better. Our study is related to both of these studies, because we use a similar within student between subject identification strategy.

Our study has some advantages over the above-mentioned studies. First of all, Aslam and Kingdon (2007) construct a dependent variable where math test scores are related to language test scores for Pakistani students. Schwerdt and Wuppermann (2009) construct a dependent variable and relate math test scores to average test scores on physics, chemistry, biology and geography, because students in the U.S. always have the same teacher for the several science subjects. Since Dutch students always have different math and physics teachers, we are able to relate test scores on math with test scores on physics. Because math compares better to physics than to language or to physics, chemistry, biology and geography together, we expect to obtain a more accurate estimate.

The second advantage is that the effect of previous lecturing styles on current student performance is likely to be smaller in this study. The studies for the U.S. and Pakistan eight grade students go to junior high (or middle school) and this serves as a bridge between elementary school and high school. While students from Pakistan also use information on eight grade students, but these students have received several years of math and science education before the TIMSS survey took place. Eight grade students in the Netherlands are in their second year of secondary school and received no more than one education year on math and physics. To test how first year lecturing styles affect student performance in the second year, we use information on whether schools assign student to classes based on their math and science skills. In this way, we test how a non-random assignment of students to classes, and hence a non-random allocation of previous lecturing styles to students, influence student performance.

We proceed as follows. Section 2 explains the identification issues and identification strategy. Section 3 discusses the data and show the descriptive statistics. In Section 4, we present and discuss the empirical results based on the within student between subject analysis. In section 5 we examine if measurement error can explain our results. In Section 6, we examine if the non-random assignment of first year students to classes can explain our results. Finally, in Section 7 we conclude.

2 Theory and Estimation Strategy

A straightforward way to examine the effect of lecturing style teaching on student performance is by estimating an education production function of the following form:

$$A_{ijk} = \beta_{0j} + S'_{ik}\beta_{1j} + T'_{ijk}\beta_{2j} + X'_{ijk}\beta_{3j} + L'_{ijk}\beta_{4j} + \epsilon_{ik}, \tag{1}$$

where A_{ijk} represents the performance of student i on subject j in school k, and where student performance depends on school (S), teacher (T) and student (X) characteristics. The effect of different lecturing styles on student performance is measured through the variable L. This variable measures the share of teaching time that teachers spend on lecturing style teaching.

If we would estimate equation (1) with an OLS estimation procedure we would ignore that there are possible selection effects. These selection effects are unobserved student, teacher and school effects that show up in the error term in a systematic manner, causing a bias in the parameter estimates. For example, schools may determine the lecturing style that is adopted by the teacher or high ability students may be assigned to high ability teachers.

To control for selection effects, we follow Aslam and Kingdon (2007) and Schwerdt and Wuppermann (2009) and identify the effect of lecturing style teaching by using a within student between subject approach. By comparing math test scores with physics test scores and by assuming that school and student characteristics influence these test scores in a similar manner, we control for selection effects at the school and the student level. Essentially, the within student between subject approach implies that

$$\begin{aligned} \Delta A_i &= (\beta_{0,m} - \beta_{0,ph}) + S'_i(\beta_{1,m} - \beta_{1,ph}) + T'_{ik,m}\beta_{2,m} - T'_{ik,ph}\beta_{2,ph} + \\ &\quad X'_i(\beta_{3,m} - \beta_{3,ph}) + L'_{i,m}\beta_{4,m} - L'_{i,ph}\beta_{4,ph} + \nu_i \end{aligned} \tag{2}$$

which is equivalent to

$$\Delta A_i = \delta + \Delta T'_i\beta_2 + \Delta L'_i\beta_4 + \nu_i, \tag{3}$$

where subscript m (ph) shows that an observation is related to math (physics) and where $\Delta A_i = A_{i,m} - A_{i,ph}$ represents the difference between math and science performance. If school and student characteristics have the same effect on student performance across subjects (and there is no reason why this would not be the case) we have that $\beta_{1,m} - \beta_{1,ph} = \beta_{3,m} - \beta_{3,ph} = 0$ so that these effects drop out of equation (3). It is however possible that certain school and student characteristics influence math performance in a different way than physics performance and therefore, as a robustness check, we should estimate equation (2) as well.

It is often mentioned that bad peers gain more by being exposed to good peers than good peers lose by being exposed to bad peers and that it is therefore crucial to control for differences in class characteristics (see, for example, Hoxby, 2000 and Lazear, 2001). However, in our sample we have identical math and physics classes and, under the assumption that peer effects affect math performance in the same way as physics performance, we automatically control for these unobserved peer effects. By adopting the within student between subject design we furthermore control for potential class size effects.

There are two problems that may occur when we estimate equation (3). The first problem is that we find that lecturing style teaching positively affects student performance but that this effect is caused by unobserved teacher characteristics. The second problem is that measurement errors in the lecturing style variables are compounded when we perform a within student between subject analysis and so it may be that we find no effect of lecturing style teaching on student performance, while in reality there is an effect. The measurement error is random noise in the lecturing style variables and the more random measurement error these variables contain the closer the estimated gradient gets to zero instead of to the true gradient. This bias due to measurement error is usually referred to as regression dilution or attenuation bias.

For this study, measurement error is problematic because the effect of measurement error in equation (3) is larger than the effect in equation (1). If the estimate in equation (3) is lower that the estimate in equation (1), we cannot identify if this difference occurs due to measurement error, or due to other factors, such as selection effects.

Below we describe how we can evaluate the selection bias due to unobserved teacher characteristics using an approach of Altonji, Elder, and Taber (2005). Consider the linear projection of ΔL onto $\Delta T'\beta_2$ and ν:

$$Proj(\Delta L|\Delta T'\beta_2, \nu) = \phi_0 + \phi_{\Delta T'\beta_2}\Delta T'\beta_2 + \phi_\nu \nu. \tag{4}$$

Clearly, the unobservable teacher characteristics are not related to differences in lecturing style if $\phi_\nu = 0$. But if $\Delta T'\beta_2$ is orthogonal to ν_i, i.e. $\phi_{\Delta T'\beta_2} = \phi_\nu$, then the part of ΔA that is related to the observables and the part related to the unobservables have the same relationship with ΔL, so that we have that:

$$\frac{Cov(\Delta T'\beta_2, \Delta L)}{Var(\Delta T'\beta_2)} = \frac{Cov(\nu, \Delta L)}{Var(\nu)} \tag{5}$$

We can now determine how large the selection on unobservables relative to the selection on unobservables would have to be in order to explain the entire lecturing style effect (β_4) given that the true effect is zero. First, we explain the difference in lecturing styles that is not caused by observed differences in teacher characteristics ($\hat{L}$). Then we replace the lecturing styles variable L by $\hat{L}$ in equation (3) and estimate this equation. By construction $\hat{L}$ is orthogonal to ΔT so that we can write the probability limit of the estimated lecturing style effect, $\hat{\beta}_4$, as:

$$plim\hat{\beta} = \beta + \frac{Cov(\nu, \Delta\hat{L})}{Var(\hat{L})}, \tag{6}$$

where

$$\frac{Cov(\nu, \Delta\hat{L})}{Var(\hat{L})} = \frac{Cov(\nu, \Delta L)}{Var(\hat{L})}, \tag{7}$$

as ΔT is orthogonal to ν_i.[1] Note, however, that it does not make sense to determine how large the selection on unobservables relative to the selection on unobservables would have to be in order to explain the entire lecturing style effect (β_4) if the lecturing style variable does not enter significantly in equation (3). This is because we would then estimate $\hat{\beta}_4 = 0$ and assume at the same time that $\beta_4 = 0$.

3 Data and descriptive statistics

We use Dutch data from the Trends in International Mathematics and Science Study 2003, better known as the TIMSS 2003 data. These data include information on the math and physics performance of students and include detailed information on students, teachers, schools and classes. The TIMSS data is a rich and unique database (especially for the Netherlands) because teachers are extensively surveyed and teacher and student information can be linked at the student level.

Although there is information on fourth and eight grade students in the TIMSS data, we use information on eight graders only. This is because fourth grade students have one teacher per year and hence we do not observe lecturing style variation over subjects (see Section 2 for the identification method). This means that we use information on 13 and 14 year old students who are in their second year of secondary school.

The sample used for the analysis contains information on 2024 students in 85 schools. Because all students in our sample are educated in math and physics by different teachers,

[1] Note that the numerator of equation (7) can be calculated using $Cov(\nu, \Delta L) = Var(\nu) \cdot Cov(\Delta T' \beta_2, \Delta L)$. The consistent estimate of β_2 is obtained by regressing ΔL on ΔT under the assumption that $\beta_4 = 0$.

there is information on 85 math teachers and 85 physics teachers. All students in the sample were tested on math and physics and the test scores are comparable when we standardize them such that the test score distribution has a mean of zero and a standard deviation of one.

Dutch students have different teachers for math, physics, chemistry, biology and geography. Because math is most similar to physics, we compare math test scores with physics test scores and math teachers with physics teachers. That math is more similar to physics than to biology, geography and chemistry, in terms of test scores, is supported by the data. The correlation coefficient between the test scores on math and the test scores on physics is slightly above 0.8.

Because it may be that different teachers choose different lecturing styles, it is important to control for differences between math and science teachers in the empirical analysis. In Table 1 we show how math and physics teachers shows are different when we compare them in gender, age, education level, years of teacher training and the minutes they give class to the students in our sample. On average, science teachers are older and higher educated are more often (and mostly) men. Math teachers are more experienced and have received more years of teacher training. With 42 minutes, the difference in teaching minutes is substantial and this may influence the effectiveness of a lecture, and on its turn student performance.

Table 1: **Descriptive Statistics: Teachers**

	Math (85 Teachers)		**Physics** (85 Teachers)		
	Mean	St.Dev.	Mean	St.Dev.	Diff.
Male Teachers	0.62	0.49	0.83	0.37	-0.21***
Age of the Teacher:					
Under 25	0.03	0.18	0.01	0.09	0.02
Between 25 and 29	0.16	0.36	0.12	0.32	0.04
Between 30 and 39	0.15	0.37	0.26	0.44	-0.11*
Between 40 and 49	0.40	0.49	0.33	0.47	0.07
Between 50 and 59	0.26	0.44	0.24	0.43	0.02
Highest Education Level:					
Secondary general/vocational education	0.04	0.19	0.04	0.19	0.00
Higher vocational	0.88	0.32	0.72	0.45	0.15***
University/PhD	0.08	0.28	0.25	0.43	-0.16***
Teaching Experience:					
Between 0 and 2 years	0.02	0.15	0.06	0.24	-0.03
Between 2 and 5 years	0.14	0.35	0.24	0.43	-0.09*
Between 5 and 10 years	0.20	0.40	0.16	0.37	0.04
More than 10 years	0.64	0.48	0.54	0.50	0.09
Years of Teacher Training	5.18	1.26	4.45	1.85	0.72***
Weekly Teaching Minutes	148.86	17.81	107.26	26.57	41.60***

Note: */**/*** means statistically significant at the 10/5/1 percent level.

Table 2 shows information on how teachers divide their time over different teaching activities and this information is used to identify the share of time that teachers spend on lecturing style teaching. The second and fourth column present the percentages of time that math and physics teachers spend to certain activities. The fifth column shows if math and physics teachers differ in the percentage of time they spend on certain activities and if this difference is significant. We note that the percentages in columns 2 and 4 both add up to 100 percent.

Lecturing style teaching is identified by the activity 'lecturing in front of the class', the second activity in column one. Physics teachers spend five percent more time on lecturing style teaching than math teachers. Reteaching activities are not considered as lecturing style teaching, although these activities can performed in front of the class. However, math and

Table 2: **Descriptive Statistics: Lecturing Styles**

	Math (85 Teachers)		Physics (85 Teachers)		
Activity	**%**	**St.Dev.**	**%**	**St.Dev.**	**Diff.**
Examine Homework	14.76	10.41	16.70	10.12	-1.94
Lecturing in front of the class	13.52	7.97	18.58	11.02	-5.05***
Students work with guidance	20.79	17.77	15.86	11.75	4.92**
Students work without guidance	26.55	20.28	14.96	14.88	11.59***
Students listen to reteaching	7.28	5.37	8.43	5.81	-1.34
Making tests or quizzes	8.50	4.85	7.85	3.90	0.64
Classroom management	4.98	5.78	6.16	5.00	-1.17*
Other activities	3.70	4.02	11.48	11.87	-7.77***

Note: */**/*** means statistically significant at the 10/5/1 percent level.

science teachers spend about the same percentage of their time to reteaching activities and so the variation of the lecturing style variable when we take first-differences does not depend on the inclusion of these reteaching activities.

To measure the effect of lecturing style teaching we construct two measures. The first measure represents the lecturing time of teachers in front of the class as a percentage of the total teaching time. However, since this measure also includes the time spent on classroom management and other activities we construct a second measure that represents the lecturing time of teachers in front of the class as a percentage of the effective teaching time. With effective teaching time we mean the time spent on all but the last two categories in Table 2.

In Table 3 we characterize the students in our sample and show the descriptive statistics on gender, ethnicity and the education level of the father and the mother. The table shows that there are about as many boys as there are girls and that, on average, fathers are higher educated than mothers. Students are labeled as immigrant student if both parents were born outside the Netherlands, and this applies to nine percent of the students in the sample. The first column shows the number of students for whom information is available. Although it is well known that parental education influences child schooling outcomes (Holmlund, Lindahl, and Plug, 2008), the table shows that information on parental education is often missing. This underlines the value of using a within student between subject identification method: even when information is missing of not available at all, the analysis controls for constant student, parental and the school-effects .

In Table 4 we show the descriptive statistics at the school level. We note that with density school area we refer to how many people live in the area where the school is. The

Table 3: **Descriptive Statistics: Students**

	Obs.	Mean	St.Dev.
Male Student	1937	0.51	0.50
Immigrant Student	2024	0.09	0.28
Highest Education Level Mother:			
Less than primary	1319	0.02	0.13
Primary	1319	0.04	0.19
Secondary general vocational	1319	0.56	0.50
Higher Vocational	1319	0.28	0.45
University - PhD	1319	0.11	0.31
Highest Education Level Father:			
Less than primary	1319	0.01	0.10
Primary	1319	0.04	0.19
Secondary general vocational	1319	0.45	0.50
Higher Vocational	1319	0.29	0.45
University - PhD	1319	0.21	0.41

Note: */**/*** statistically significant at the 10/5/1 percent level.

table shows that schools vary substantially with respect to the number of people who live in the area where the school is, the percentage of disadvantage students, class size and the job duration of the principal. The job duration of the principal is very much skewed, with many principals only having a few years of experience. The reason why we include job duration of the principal as one of the control variables is because this variable may be related to the probability that new education programmes are implemented or may be related to the past performance of the school.

4 Empirical Results

In this section we present the estimation results based on the formulated specifications in Section 2. We first present the estimation results when we estimate the model presented by equation (1) seperately for math and physics. The estimation results related to math are presented in Table 5. We perform four alternative estimations. For alternatives one and two we use the first definition of lecturing style teaching and for alternatives three and four we use the second definition of lecturing style teaching. For each of the different definitions of lecturing style teaching we perform one regression where we only control for teacher and student characteristics, and one regression where we also control for school and

Table 4: **Descriptive Statistics: School and Class**

	Mean	**St.Dev.**
School Characteristics (N = 85)		
Living Density School Area:		
Greater than 500000	0.05	0.22
Between 100001 and 500000	0.24	0.43
Between 50001 and 100000	0.21	0.41
Between 15001 and 50000	0.44	0.50
Between 3001 and 15000	0.05	0.22
Percentage Disadvantage Students:		
Less than 10 percent	0.65	0.48
Between 11 and 25 percent	0.24	0.43
Between 26 and 50 percent	0.08	0.27
More than 50 percent	0.02	0.16
Job Duration Principal (years)	8.14	7.59
Class size	23.81	3.91

class characteristics. The control variables that we use are discussed in Section 3. We note that, although we did not show the averages number of boys and immigrant students per class in Section 3, we do include these variables as control variables in the analysis.

The lecturing style variable is positive and significant in all specifications, suggesting more lecturing style teaching would improve the math performance of students. We note that lecturing style observations are fractions and not percentages. The lecturing style coefficient of the first alternative therefore means that an increase of lecturing style teaching of one percent is associated with a test score on math that is 2.6 percent of a standard deviation higher. An interesting result is that the estimates do not support the idea that a more personal lecturing approach influences student performance positively, as is often believed in the Netherlands. While the estimation results show that lecturing style teaching affects the math performance of students positively, they at the same time show that the results are sensitive to the lecturing style definition and tend to be lower when we control for school- and class-effects.

With respect to the teacher variables that are included in the regression, we find that teachers with a university degree or higher, teachers aged between 40 and 49 and teachers with more experience perform better. Teachers who have received more teacher training perform less well, but it is unlikely that this is caused by the teacher training itself. Probably, those teachers who performed less well are also those teachers who receive relatively more teacher training. When teachers teaches more minutes this increases the math performance of students and this is consistent with the idea that (more) teaching positively affects student performance.

In Table 6 we present the estimation results related to physics in the same way as we did for math. We find that lecturing style teaching is positively significant if we only control for student and teacher effects, but that this significance disappears when we also control for school and class effects. Students with higher educated and more experience teachers tend to perform better. Again, teachers who receive more years of teacher training perform less well, but this does not represent the causal relationship between teacher training and student performance but merely shows that teachers who perform less well receive more training.

The estimates presented in Tables 5 and 6 ignore the potential selection effects. Therefore all parameter estimates are biased if unobservable student, teacher and school effects show up in the error term in a systematic manner. Schools, for example, may determine the lecturing style that is adopted by the teacher and these schools may also attract students who are perform on average better (or worse). Also high ability students may be linked to high ability teachers within schools, or high ability teachers and high ability students may be more likely to self-select in the same (type of) school.

In Table 7 we report the within student between subject estimation results and this estimation model is represented by equation (3). The dependent variable is the within student difference in standardized test scores between math and physics and the explanatory variables are the teacher characteristics that are also presented in Table 5 and 6.

The table shows that differences between math and science test scores are not explained by differences lecturing style teaching. While Schwerdt and Wuppermann (2009) find that students perform somewhat better if teachers lecture more in front of the class, our results do not confirm this finding for the Netherlands. However, the empirical results of this study are in line with those of Schwerdt and Wuppermann (2009), in the sense that both studies do not confirm that a more personal teaching style would increase the performance of students.

When we compare the estimates of Table 7 with those of Table 5/6 it is striking to see that most of the teacher characteristics are not related to the difference in math and physics performance, even though this result is consistent with some of the literature that measures the relation between teacher characteristics on student performance (see for example Hanushek, Kain, and Rivkin, 2005, Hanushek and Rivkin, 2006 and Aaronson, Barrow, and Sander, 2007) . We emphasize, that the teacher differences presented in Table 7 are difficult to interpreted. However, we deliberately defined the variables in the within student between subject model in the same way as in the models underlying Table 5 and 6. In this way the

Table 5: OLS Estimates for Math

	(1)		(2)		(3)		(4)	
Lecturing style [def. 1]	2.556***	(0.254)	1.035***	(0.249)				
Lecturing style [def. 2]					1.591***	(0.166)	1.023***	(0.156)
Male teacher	0.104**	(0.049)	0.026	(0.049)	0.138***	(0.049)	0.047	(0.049)
Teachers' age:								
Under 25	0.274*	(0.167)	0.271*	(0.163)	0.395*	(0.167)	0.347*	(0.162)
Between 25 and 29	-0.463**	(0.105)	-0.443***	(0.100)	-0.448**	(0.105)	-0.435***	(0.099)
Between 30 and 39	0.031	(0.069)	0.112*	(0.065)	0.037	(0.069)	0.123*	(0.065)
Between 50 and 59	-0.045***	(0.056)	-0.280***	(0.055)	-0.403***	(0.055)	-0.258***	(0.054)
Teachers' Education:								
Secondary general/vocational education	-0.179	(0.125)	-0.167	(0.117)	-0.217	(0.126)	-0.175	(0.116)
University/PhD	0.083	(0.089)	0.180**	(0.089)	0.122	(0.089)	0.163**	(0.088)
Teachers' Experience:								
Between 0 and 2 years	0.355**	(0.156)	-0.110	(0.147)	0.445**	(0.156)	-0.100	(0.146)
Between 2 and 5 years	-0.235**	(0.103)	-0.290***	(0.097)	-0.282**	(0.103)	-0.320***	(0.096)
Between 5 and 10 years	-0.263***	(0.064)	-0.416***	(0.060)	-0.248***	(0.064)	-0.403***	(0.060)
Years Teacher Training	-0.140***	(0.021)	-0.075***	(0.021)	-0.130***	(0.021)	-0.073***	(0.020)
Weekly Teaching Minutes	0.003**	(0.001)	0.003**	(0.001)	0.002**	(0.001)	0.002**	(0.001)
Constant	0.054	(0.226)	-1.153***	(0.268)	0.111	(0.226)	-1.271***	(0.267)
Student effect	Yes		Yes		Yes		Yes	
Class effects	No		Yes		No		Yes	
School effects	No		Yes		No		Yes	
Adjusted R^2	0.174		0.303		0.171		0.312	
Number of observations	2024		2024		2024		2024	

Note: Teachers reference group are women with a higher vocational education level, aged between 40-49 with more than 10 years of teaching experience. */**/** means statistically significant at the 10/5/1 percent level. Clustered standard errors (school level) are printed in parentheses.

Table 6: OLS Estimates for Physics

	(1)		(2)		(3)		(4)	
Lecturing style [def. 1]	0.688***	(0.182)	0.041	(0.177)				
Lecturing style [def. 2]					0.444***	(0.145)	-0.023	(0.140)
Male teacher	-0.202***	(0.057)	-0.132**	(0.058)	-0.189***	(0.057)	-0.127**	(0.058)
Teachers' age:								
Under 25	-0.524**	(0.265)	-0.106	(0.252)	-0.529**	(0.265)	-0.105	(0.252)
Between 25 and 29	0.347***	(0.081)	0.407***	(0.078)	0.350***	(0.081)	0.409***	(0.078)
Between 30 and 39	-0.111*	(0.063)	0.010	(0.063)	-0.103*	(0.063)	0.012	(0.063)
Between 50 and 59	-0.094*	(0.053)	-0.229***	(0.052)	-0.090*	(0.053)	-0.229***	(0.052)
Teachers' Education:								
Secondary general/vocational education	-0.486***	(0.146)	-0.254*	(0.139)	-0.488***	(0.146)	-0.252*	(0.139)
University/PhD	0.527***	(0.058)	0.521***	(0.057)	0.531***	(0.058)	0.521***	(0.057)
Teachers' Experience:								
Between 0 and 2 years	-0.218*	(0.112)	-0.232**	(0.108)	-0.231**	(0.112)	-0.233**	(0.108)
Between 2 and 5 years	-0.108*	(0.065)	-0.225***	(0.063)	-0.102	(0.066)	-0.222***	(0.063)
Between 5 and 10 years	0.165**	(0.074)	-0.012	(0.075)	0.189**	(0.074)	-0.012	(0.075)
Years Teacher Training	-0.090***	(0.014)	-0.049***	(0.001)	-0.094***	(0.014)	-0.049***	(0.001)
Weekly Teaching Minutes	-0.001	(0.001)	0.001	(0.001)	-0.001	(0.001)	0.001	(0.001)
Constant	0.352**	(0.135)	-0.974***	(0.216)	0.330**	(0.136)	-0.971***	(0.216)
Student effect	Yes		Yes		Yes		Yes	
Class effects	No		Yes		No		Yes	
School effects	No		Yes		No		Yes	
Adjusted R^2	0.269		0.357		0.268		0.357	
Number of observations	2024		2024		2024		2024	

Note: teachers reference group are women with a higher vocational education level, aged between 40-49 with more than 10 years of teaching experience. */**/** means statistically significant at the 10/5/1 percent level. Clustered standard errors (school level) are printed in parentheses.

conditions under which one model transforms into another are clear.

We note, however, that we also performed an analysis with redefined teacher variables. In this analysis we used dummy variables that represent the difference in age intervals, the difference in the number of education levels and included a variable that captured the difference in years of experience. The estimates of this analysis show that most of the teacher characteristics are not related to the difference in math and physics performance, with exception of the difference in education level of teachers . When math teachers are educated one level more (less), then this has a small but positive (negative) effect on the test score difference between math and science and this result is consistent with Clotfelter, Ladd, and Vigdor (2006) and Clotfelter, Ladd, and Vigdor (2007).

Furthermore we re-estimated equation (3) and included the school and student characteristics. As we discussed in Section 3, we thereby test if the lecturing style estimate changes if we do not assume that school and student characteristics have the same effect on math and physics performance. The empirical findings remain similar to those reported in Table 7.

5 Can measurement error explain the results?

When we perform a within student between subject regression, the measurement error in the lecturing style variables are compounded and so the effect of measurement error in equation (3) is larger than in equation (1). The consequence of measurement error is that the estimated gradient tends to go to zero instead of to the true gradient and this is generally referred to as attenuation bias. For this study, this can be problematic since we cannot identify whether the difference between the lecturing style estimates in (1) and (3) are due to measurement error or are due to other factors, such as selection effects. In this section we examine if the difference between the lecturing style estimates is the result of measurement error by taking first-differences.

We demonstrate how measurement error can drive the lecturing style estimates towards zero and then determine if the extra measurement error that is caused by taking first-differences can explain the difference in lecturing style estimates between equations (1) and (3). For this purpose, we assume that measurement errors on the explanatory variables are independent (see Cragg, 1994) and classical, so that the effects of measurement error can be examined as if student performance depends solely on lecturing style teaching.

In the first step we estimate a pooled panel model using OLS and determine the true

Table 7: Within Student Between Subject Estimates

Differences in:	**(1)**		**(2)**	
Lecturing style [def. 1]	-0.049	(0.128)		
Lecturing style [def. 2]			-0.017	(0.096)
Male teacher	-0.004	(0.030)	-0.005	(0.031)
Teachers' age:				
Under 25	-0.063	(0.092)	-0.061	(0.093)
Between 25 and 29	-0.154***	(0.045)	-0.153***	(0.045)
Between 30 and 39	-0.074***	(0.036)	-0.074***	(0.036)
Between 50 and 59	-0.083***	(0.030)	-0.084***	(0.030)
Teachers' Education:				
Secondary general/vocational education	0.128*	(0.066)	0.126*	(0.067)
University/PhD	-0.067*	(0.038)	-0.067*	(0.038)
Teachers' Experience:				
Between 0 and 2 years	0.039	(0.066)	0.040	(0.067)
Between 2 and 5 years	0.028	(0.044)	0.028	(0.045)
Between 5 and 10 years	0.014	(0.034)	0.013	(0.034)
Years Teacher Training	0.010	(0.008)	0.010	(0.009)
Weekly Teaching Minutes	0.001	(0.001)	0.001	(0.001)
Constant	-0.040	(0.028)	-0.038	(0.027)
Adjusted R^2	0.09		0.09	
Number of observations	2024		2024	

Note: Teachers reference group are women with a higher vocational education level, aged between 40-49 with more than 10 years of teaching experience. */**/** means statistically significant at the 10/5/1 percent level. Clustered standard errors (school level) in parentheses.

lecturing style estimate for different amounts of measurement error. In the second step, we determine what correlation coefficient between the lecturing style variables is needed to cause a drop from the true parameter value to the parameter estimate reported in Table 7. In this way we examine the effect of the extra measurement error by taking first-differences. In addition, we simulate the lecturing style parameter for different levels of measurement error and correlation values.

We start with considering a pooled panel model where student performance, A, depends on the lecturing style variable :

$$A_{is} = \gamma L^*_{is} + \eta_{is}, \tag{8}$$

and where A_{is} represents student performance of student i in subject s that represents mathematics and physics. We assume that the observed lecturing style, L_{is}, is measured with error and that it depends on the true lecturing style with an additive random measurement error, i.e. $L_{is} = L^*_{is} + \omega_{is}$. If we would estimate equation (8), where the true lecturing style variable, L^*_{is}, is replaced by the observed lecturing style variable, L_{is} then[2]

$$\mathrm{p}\lim \widehat{\gamma}_{pooled} = \gamma \cdot (1 - \frac{\sigma^2_\omega}{\sigma^2_{L^*} + \sigma^2_\omega}). \tag{9}$$

The second term between parentheses in equation (9) represents the bias due to measurement error. We estimated this pooled panel equation, i.e. equation (8), using ordinary least squares with clustured standard errors (student level), and found a positive and significant lecturing style estimate, $\widehat{\gamma}_{pooled}$, of 0.39. The full estimation results of the pooled OLS are presented in Appendix A. Using $\widehat{\gamma}_{pooled}$ we can calculate the true parameter, γ, for different hypothetical amounts of measurement error and these true parameters are reported in Table 8.

The first column shows the percentage of measurement error (ME), and the second column shows the true parameter estimate that is associated with this hypothetical measurement error. Clearly, the true parameter is equal to the estimated parameter in the absense of measurement error. If there is a positive amount of measurement error, the true parameter is more positive than the estimate obtained from the pooled OLS, and so the pooled OLS can be viewed as a lower bound estimate.

Column three shows the variance of the lecturing style variable used in the pooled OLS and the variances in column four and five are calculated using this variance and the percentage of measurement error. If, for example, the measurement error is ten percent then σ^2_ω

[2]See Cameron and Trivedi (2005)

Table 8: True lecturing style parameters conditional on measurement error

% of ME	γ	σ^2_L	σ^2_w	$\sigma^2_{L^*}$
0 %	0.394	0.102	0.000	0.102
5 %	0.415	0.102	0.005	0.097
10 %	0.438	0.102	0.010	0.092
20 %	0.493	0.102	0.020	0.082
40 %	0.657	0.102	0.041	0.061
50 %	0.788	0.102	0.051	0.051
70 %	1.313	0.102	0.071	0.031
90 %	3.944	0.102	0.092	0.01

equals 0.0102, which is 10 percent of σ^2_L. The variances in column three up to five are used later to show if additional measurement error in the first-difference analysis can explain the change in estimate from $\widehat{\gamma}_{pooled}$ to the first-difference estimate, $\widehat{\gamma}_{FD}$.[3]

In the within student between subject analysis we recognize that regressor L is positively related, rather than independent over s for given i and the analysis is similar to a first-difference analysis:

$$\begin{aligned} \Delta A_i =& \gamma \Delta L^*_i + \Delta\eta_i \\ =& \gamma \Delta L_i + \Delta\eta_i - \gamma\omega_i. \end{aligned} \tag{10}$$

If we define $\rho =$Cor$[L^*_m, L^*_{ph}]$, then

$$\text{p}\lim \widehat{\gamma}_{FD} = \gamma \cdot (1 - \frac{\sigma^2_\omega}{(1-\rho)\sigma^2_{L^*_i} + \sigma^2_\omega}), \tag{11}$$

using $V[\Delta\omega_i]$=$2V[\omega_i]$ and $V[\Delta L^*_i]$=$2(1-\rho)V[L^*_i]$ (see Cameron and Trivedi, 2005, p. 905) and where subscript FD stands for first-difference. When $\rho = 0$, the measurement error of the first-difference analysis resembles that of pooled OLS analysis. For our sample, we observe Cor$[L_{im}, L_{iph}]$= 0.243 and in the presence of measurement error this correlation coefficient is a lower bound. The true correlation coefficient, $\rho =$Cor$[L^*_m, L^*_{ph}]$, must therefore be higher. The relevant question, for this study is whether the extra measurement error due to the true correlation between the lecturing styles variables can be high enough to drive the

[3]In this section we only present the results using the first lecturing style definition (see Section 2). The conclusions when using the second lecturing style definition led to similar conclusions. The results are available on request.

Table 9: Simulation of true correlation between the lecturing style variables

% ME_{pooled}	γ	$\widehat{\gamma}_{FD}$	ρ
0 %	0.394	-0.049	0.000
5 %	0.415	-0.049	1.006
10 %	0.438	-0.049	1.011
20 %	0.493	-0.049	1.022
40 %	0.657	-0.049	1.046
50 %	0.788	-0.049	1.058
70 %	1.313	-0.049	1.084
90 %	3.944	-0.049	1.111

estimate of the pooled OLS towards the estimate of the first-difference analysis.

In Table 8, we showed the true lecturing style parameters given the hypothetical amounts of measurement error. On the basis of these true parameter values and equation (11) we simulate the true correlation between the lecturing style variables given that the parameter drops from γ to $\widehat{\gamma}_{FD}$. These calculated true correlation coefficients, ρ, are shown in Table 9.

The first two columns of Table 9 show the hypothetical measurement error with the associated true parameter from Table 8. Column three and four show the first-difference parameter and the simulated correlation coefficient that is needed to explain the drop in the estimation parameter from γ to $\widehat{\gamma}_{FD}$. The table clearly shows that there are no feasible values for ρ that can explain the drop in the estimation parameter, because the predicted values for ρ are higher than one, while its value must lie between zero and one. The reason why values are predicted that are slightly above one, is because the first-difference estimate is negative. If the estimate would have been slightly higher than zero, we would have found ρ-values that are slightly lower than one. Hence we conclude that the drop in the estimation parameter from γ to $\widehat{\gamma}_{FD}$ cannot be due to measurement error.

However, the change in the estimation parameter may be caused by measurement error *and* other factors, such as the fact that we control for selection effects in the first-difference analysis. This means that if we would isolate only the measurement error, there may be feasible values of ρ that cause a drop in the first-difference analysis from γ to a different value than $\widehat{\gamma}_{FD}$, say $\widetilde{\gamma}$, that lies between $\widehat{\gamma}_{FD}$ and $\widehat{\gamma}_{pooled}$. To have an idea of how $\widetilde{\gamma}$ is influenced by measurement error only, we simulate it for different values of ρ, hypothetical measurement errors and γ and the results are shown in Table 10.

The first column of Table 10 shows the percentage of hypothetical measurement error in the pooled OLS analysisis. The upper panel of Table 10 shows the simulated values of $\widetilde{\gamma}$ for different values of ρ, while the lower panel shows the simulated measurement error in the

Table 10: Simulation of $\widetilde{\gamma}$ and the first-difference measurement error

	Simulated value of $\widetilde{\gamma}$ given:				
% ME_{pooled}	ρ=0	ρ=0.243*	ρ=0.5	ρ=0.75	ρ=0.90
0 %	0.394	0.394	0.394	0.394	0.394
5 %	0.394	0.388	0.375	0.343	0.272
10 %	0.394	0.382	0.358	0.303	0.207
20 %	0.394	0.370	0.328	0.246	0.141
40 %	0.394	0.349	0.281	0.179	0.086
50 %	0.394	0.340	0.263	0.158	0.072
70 %	0.394	0.322	0.232	0.127	0.054
90 %	0.394	0.306	0.207	0.106	0.043
	Simulated ME_{FD} given:				
% ME_{pooled}	ρ=0	ρ=0.243*	ρ=0.5	ρ=0.75	ρ=0.90
0 %	0.00	0.00	0.00	0.00	0.00
5 %	5.00	6.50	9.52	17.39	34.48
10 %	10.00	12.80	18.18	30.77	52.63
20 %	20.00	24.83	33.33	50.00	71.43
40 %	40.00	46.83	57.14	72.73	86.96
50 %	50.00	56.92	66.67	80.00	90.91
70 %	70.00	75.50	82.35	90.32	95.89
90 %	90.00	92.24	94.74	97.30	98.90

*This correlation resembles the correlation between L_m and L_{ph} that is observed in the data.

first-difference analysis for different values of ρ.

When we concentrate on the upper panel of Table 10, the first row shows that $\widetilde{\gamma}$ is similar to the pooled OLS estimator, for all values of ρ, if there is no measurement errror. This is intuitive, because there is no measurement error in the pooled OLS analysis, and so there is no measurement error that can be magnified in the first-difference analysis. The same holds when we consider the values of $\widetilde{\gamma}$ for $\rho = 0$, and this is because measurement error is magnified only if $\mathrm{Cor}[L^*_m, L^*_{ph}] > 0$. On the basis of the upper panel we conclude, that the simulated values of $\widetilde{\gamma}$ are substantially larger than zero for all ρ up to 0.75, irrespective, of the amount of measurement error.

The lower panel of Table 10 shows the amount of measurement error in the first-difference analysis when we simulate values of $\widetilde{\gamma}$ for different ρ's and amounts of measurement error. If $ME_{pooled} > 0$, we find that an increase of ρ increases the measurement error in the first-difference analysis compared to the pooled OLS analysis and this increase can be substantial. For example, if ρ is 0.5 and the measurement error in the pooled OLS is 50 percent, then we find that the measurement error in the first-difference analysis increases to almost 67 percent, which is an increase of 17 percent. However, despite this 17 percent extra measurement error we would still find an estimate of 0.263, which is substantially greater than zero.

We conclude that the extra measurement error in the first-difference analysis drives the estimation coefficient of the lecturing style variable (substantially) towards zero, but the results above clearly show that it is unlikely that the estimated parameter will be close to zero. In fact, even in the presence of a substantial amount of measurement error together with a correlation between L^*_m and L^*_{ph} that may go up to 0.75, it is more likely that we find a positive lecturing style estimate.

6 Can the non-random assignment of first year students to classes explain the results?

Dutch students who are in their second year of secondary school in 2003 may be assigned to teachers in the first year of secondary school based on the math and physics skills. Therefore the non-observed teacher characteristics and lecturing styles of the first year of secondary school may influence the lecturing style estimate.

For several reasons it is unlikely that schools in the Netherlands assign their first year students to classes based on their math and science skill. First of all, students start with math and physics in their first year of secondary school and so schools cannot observe their math and science ability at the beginning of this year. Second, primary schools give an advice on the level of secondary school they think a student should go to. In general, this advice is consistent with the secondary school levels students go to and this means that the study advice of primary schools is a rather uninformative measure of the learning abilities of students. However, it is possible that first year students are assigned to classes on the basis of test score of a national test that each Dutch student makes in the last grade of their primary school (the CITO test). Therefore we cannot exclude the possibility that first year students are assigned to classes on the basis of their test score, and the non-random allocation of students to classes may affect the lecturing style estimate.

In the data there is information on whether schools assign students to classes on the basis of the math and science skills, and we will use this information to test how the non-random assignment of students to classes, and hence a non-random allocation of previous lecturing styles to students, influence student performance. When schools allocate students on the basis of their math and science skills we assume that there is a positive probability that this happens in the first year of secondary school. When schools do not allocate students on the basis of their math or science skills we assume that there is no assignment of students to classes on the basis of their math or science skills. We distinguish between weak sorting and strong sorting. Strong sorting means that students are assigned to classes based on their science *and* math skills, and weak sorting means that students are assigned to classes based on their math skills or based on their physics skills, but not both.

Not all schools reply to the sorting question, and therefore we perform three within student between subject regressions for both lecturing style definitions. In the first regression we estimate equation (3) for the schools who answered the sorting questions. In this way we examine how the baseline estimates change because the sample changes. In the second regression we estimate equation (3), include a strong sorting dummy and interact this dummy with the lecturing style variable. The third regression is similar to the second regression, but now we use the weak sorting dummy instead of the strong sorting dummy. In this section we are only interested in how the lecturing style estimate changes if we include the sorting dummies and so we do not report the estimation results on the other explanatory variables, although we note that they are very similar to the ones reported in Table 7. The estimation results are shown in Table 11.

The baseline estimates are very similar to those reported in Table 7. The lecturing style variable and the lecturing style variable interacted with the sorting dummies are always insignificant. Also the separate sorting dummies are insignificant, indicating that the assign-

Table 11: Estimates controlling for non-random assignment of students to classes.

	Lecturing Styles Definition 1		**Lecturing Styles Definition 2**	
Baseline Estimation:				
Lecturing style	-0.088	(0.132)	-0.056	(0.099)
Estimation Strong Sorting:				
Lecturing style	-0.096	(0.150)	-0.074	(0.107)
Strong sorting	0.427	(0.060)	0.051	(0.059)
Interaction	0.101	(0.340)	0.236	(0.370)
Estimation Weak Sorting:				
Lecturing style	-0.042	(0.156)	-0.084	(0.156)
Weak sorting	-0.001	(0.048)	0.025	(0.048)
Interaction	-0.156	(0.297)	0.209	(0.300)
Teacher effects	Yes		Yes	
Adj. R^2 - Baseline	0.009		0.009	
Adj. R^2 - Strong Sorting	0.008		0.008	
Adj. R^2 - Weak Sorting	0.008		0.008	
Number of observations	1921		1921	

Note II: */**/** means statistically significant at the 10/5/1 percent level. Clustered standard errors (school level) are printed in parenthesis.

ment rule of schools do not not influence the test score differences of students, at least not in their second year of secondary education.

7 Concluding remarks

This study examines whether the share of time that teachers spend on lecturing style teaching influences the cognitive performance of Dutch students. Dutch teachers tend to give less lectures in front of the class, and instead ‘choose’ for a more personal approach, because it is believed that this positively affects student performance. However, the downside of a more personal approach is that it is time intensive and possibly eliminates the complementary effects of lecturing style teaching. A more personal teaching style may therefore be a less efficient teaching style. By focusing more on the things that teachers do in class, rather than on the characteristics they possess, we obtain insights on whether lecturing in front of the class is old fashioned and whether a more personal lecturing style is beneficial for the cognitive performance of students.

To identify the effect of lecturing style teaching on student performance, we perform a within student between subject approach. In this approach we examine how test score differences on math and physics are explained by differences in the share of time that math and pysics teachers spend on lecturing style teaching, while controlling for teacher, school, student effects. By comparing test scores on math with those on physics for each student and by assuming that school and student characteristics influence these test scores in the same way, we control for selection effects at the school and student level.

We find no relationship between lecturing style teaching and student performance. Hence, our results do not support the idea that lecturing in front of the class is old fashioned or that a more personal teaching style would be more beneficial for the cognitive performance of students.

In the within student between subject approach we consider the difference in time that math and physics teachers spend on lecturing style teaching. By taking differences the measurement errors in the lecturing style variables are compounded and, therefore, we examine how likely it is that measurement error is responsible for finding an estimate close to zero. We find that it is unlikely that measurement error prevents us to find a positive lecturing style estimate. In any case, the first-difference estimate represents a lower bound and so we reject the hypothesis that students perform less well if teachers perform relatively more time in front of the class, with or without measurement error.

Dutch students who are in their second year of secondary school in 2003 may be non-randomly assigned to classes in the first year of secondary school and so teacher characteristics and lecturing styles of the first year of secondary school may influence the lecturing style estimtate. Using information on whether schools assign students to classes based on their math or science skills, we find that the lecturing style estimate is not influenced by this non-random assignment of students to classes. Furthermore, we find no evidence that the non-random assignment of students to classes influence student performance.

Apparently, student performance and lecturing style teaching are not related. It may be that we find no effect because the time spent on certain teaching activities is less important than the timing of these activities. In the Netherlands, for example, it was common that teachers lectured in front of the class for the entire lesson. These teachers were still able to observe whether a student understood a certain topic, based on the tests they gave and on their own observation/intuition. So even though the common lecturing style was lecturing style teaching, this did not prevent teachers to give students who needed it the most extra attention. It is, however, important to recognize that the explanation given above is only consistent with our empirical findings when teachers time their teaching activities optimally, or when teachers do not time their teaching activities optimally, but that non-optimal timing is not structurally related with how math and physics teachers allocate their time to different teaching activities.

Appendix A

In this Appendix we show the estimation results of the pooled OLS. Alternative one shows the results when we control for student and teacher effects only, and alternative two shows the results when we also control for school and class effects. In Section 5 we use the lecturing style estimate of Alternative (2), because this alternative gives the lowest lecturing style estimate and it is more likely that measurement error will drive this estimate to a value close to zero.

Table 12: Pooled OLS estimates

	(1)		(2)	
Lecturing style [def. 1]	1.246***	(0.146)	0.394***	(0.140)
Male teacher	-0.067	(0.035)	-0.007	(0.034)
Teachers' age:				
Under 25	0.291***	(0.119)	0.391***	(0.112)
Between 25 and 29	0.000	(0.060)	-0.008	(0.056)
Between 30 and 39	-0.045	(0.047)	0.084**	(0.045)
Between 50 and 59	-0.286***	(0.039)	-0.305***	(0.037)
Teachers' Education:				
Secondary general/vocational education	-0.586***	(0.091)	-0.366***	(0.085)
University/PhD	0.397***	(0.049)	0.413***	(0.047)
Teachers' Experience:				
Between 0 and 2 years	-0.305***	(0.087)	-0.361***	(0.081)
Between 2 and 5 years	-0.250***	(0.052)	-0.299***	(0.048)
Between 5 and 10 years	-0.132***	(0.049)	-0.276***	(0.046)
Years Teacher Training	-0.084***	(0.011)	-0.040***	(0.011)
Weekly Teaching Minutes	0.002***	(0.001)	0.001***	(0.000)
Constant	0.073***	(0.106)	-1.182***	(0.154)
Student effect	Yes		Yes	
Class effects	No		Yes	
School effects	No		Yes	
Adjusted R^2	0.298		0.113	
Number of observations	4048		4048	

Note: Teachers reference group are women with a higher vocational education level, aged between 40-49 with more than 10 years of teaching experience. */**/*** means statistically significant at the 10/5/1 percent level. Clustered standard errors (student level) are printed in parentheses.

References

Daniel Aaronson, Lisa Barrow, and William Sander. Teachers and student achievement in the chicago public high schools. *Journal of Labor Economics*, 25:95–135, 2007.

Joseph Altonji, Todd E. Elder, and Christopher R. Taber. Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy*, 113(1):151–184, February 2005.

M. Aslam and G. Kingdon. What can teachers do to raise pupil achievement. working paper series, The centre for the study of african economies, 2007.

A.C. Cameron and P.K. Trivedi. *Microeconometrics: methods and applications*. Cambridge University Press, New Yor, 2005.

Charles T. Clotfelter, Helen F. Ladd, and Jacob L. Vigdor. Teacher-student matching and the assessment of teacher effectiveness. NBER Working Papers 11936, National Bureau of Economic Research, Inc, 2006.

Charles T. Clotfelter, Helen F. Ladd, and Jacob L. Vigdor. Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects. NBER Working Papers 13617, National Bureau of Economic Research, Inc, 2007.

J.G. Cragg. Making good inferences from bad data. *Canadian Journal of Economics*, 27: 776–800, 1994.

Eric Hanushek, John F. Kain, and Steven G. Rivkin. Teachers, schools, and academic achievement. *Econometrica*, 73(2):417–458, 2005.

Eric A. Hanushek and Steven G. Rivkin. *Teacher Quality*, volume 2 of *Handbook of the Economics of Education*, chapter 18, pages 1051–1078. Elsevier, 2006.

Helena Holmlund, Mikael Lindahl, and Erik Plug. The causal effect of parent's schooling on children's schooling: A comparison of estimation methods. IZA Discussion Papers 3630, Institute for the Study of Labor (IZA), 2008.

Caroline Hoxby. Peer effects in the classroom: Learning from gender and race variation. NBER Working Papers 7867, National Bureau of Economic Research, Inc, August 2000.

Edward P. Lazear. Educational production. *The Quarterly Journal of Economics*, 116(3): 777–803, 2001.

Jesse Rothstein. Teacher quality in educational production: Tracking, decay, and student achievement. NBER Working Papers 14442, National Bureau of Economic Research, Inc, 2008.

Guido Schwerdt and Amelie C. Wuppermann. Is traditional teaching all that bad? a within-student between subject approach. *working paper*, 2009.

TIER WORKING PAPER SERIES
TIER WP 10/16

ISBN 978-94-003-0015-6