ifo WORKING PAPERS

243
2017

December 2017

The Effect of Pre-Service Cognitive and Pedagogical Teacher Skills on Student Achievement Gains:

Evidence from German Entry Screening Exams

Bernhard Enzi

ifo INSTITUTE

Leibniz Institute for Economic Research
at the University of Munich

Impressum:

ifo Working Papers
Publisher and distributor: ifo Institute – Leibniz Institute for Economic Research at the University of Munich
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49(0)89 9224 0, Telefax +49(0)89 985369, email ifo@ifo.de
www.cesifo-group.de

An electronic version of the paper may be downloaded from the ifo website:
www.cesifo-group.de

The Effect of Pre-Service Cognitive and Pedagogical Teacher Skills on Student Achievement Gains: Evidence from German Entry Screening Exams

Bernhard Enzi †

Abstract

Information about teachers' effectiveness at the hiring stage is particularly scarce despite its importance for personnel decisions. Using the German setting of teacher training, I investigate the relationship of teachers' pre-service cognitive and pedagogical skills as measured by two state examinations and the high-school GPA on later effectiveness. I apply standard value-added models to rich German student-achievement panel data and find that being in the top quartile in these skill domains is linked with significantly higher teacher effectiveness. Better teacher skills are associated with a more efficient way of classroom management.

Keywords: Teacher, value-added, cognitive skills, student achievement.

JEL classifications: I21, J24, J45, H75.

December 4, 2017

† ifo Institute – Leibniz Institute for Economic Research at the University of Munich. enzi@ifo.de.

1. Introduction

How can teachers be selected more efficiently? Teacher quality research shows that there are few determinants that help identify effective teachers in-service and even fewer before-service. This study increases the knowledge in this area by investigating the screening devices employed by German education policy to determine entry into the teaching profession: the high school grade point average (GPA) and the grades received in two state examinations.

I provide the first empirical evidence on how the screening process for teachers in the German education system is related to teacher effectiveness. Screening is a two-step procedure: Entry into field of study is determined by the high school GPA and entry into profession is primarily based on GPAs from two state examinations which are assigned by the respective federal state's education ministry during the teacher training program.

The first examination measures cognitive and theoretical pedagogical skills; the second examination practical pedagogical skills based, partially, on demonstration lessons graded by a head teacher. The high school GPA (or "Abitur" grade) can be viewed as a second-order screening device that partially determines entry into the teacher training programs at the university level. Effective selection of the teaching force is likely to be a cost-effective way to increase educational outcomes as improving student-teacher matches will lead to real educational and accompanying economic gains. Furthermore, as dismissing a teacher can be costly or even judicially virtually impossible, effective screening is all the more important.

I apply a standard value-added model to very rich micro-level data from the German National Educational Panel Study (NEPS) that comprises a nationally representative sample of secondary school students to analyze the determinants of teacher quality. I find that being in the top quartile of the three skill measures is associated with higher teacher effectiveness in mathematics. However, no significant effects exist for language teachers. Better teacher skills are associated with a more efficient way of classroom management.

This study is organized as follows: Section 2 describes the literature on determinants of teacher quality and the institutional background regarding teacher training and the teacher labor market in Germany. The data are presented in Section 3. Section 4 presents the value-added framework and the strategy for estimating the parameters of interest. Main results for the screening variables are presented in Section 5 and their in-class time activity correlates in Section 6. Section 7 contains results for different student subgroups and nonlinear effects. Section 8 concludes.

2. Background

2.1 Literature on Teacher Effectiveness

There is a great deal of work on teacher effectiveness in the economics of education literature. For a general overview, see Hanushek and Rivkin (2006). This literature can be divided into three subfields: estimation of teacher quality variance, investigations of determinants of quality differences, and analyses of teacher policies or interventions (Chetty, Friedman, and Rockoff 2014).

The strand of the literature that focuses on the effects of teacher education or certification is the one most closely related to this study. In the United States, this research is mainly motivated by the fact that many states provide additional compensation to teachers who hold a master's degree or have been certified by the National Board of Professional Teaching Standards. However, there appears to be little or even no impact of an advanced degree on student learning (Clotfelter, Ladd, and Vigdor 2006; Goldhaber 2002; Harris and Sass 2011; Rivkin, Hanushek, and Kain 2005), and the literature provides mixed evidence as to the effects of board certification. Some studies such as Cantrell et al. (2008) for Los Angeles and Ladd, Sass, and Harris (2007) for North Carolina identify a positive correlation between effectiveness and NBPTS certification. Goldhaber and Anthony (2007), however, show that NBTPS certification does not enhance teacher effectiveness. Moreover, any observed correlation between certification and effectiveness might be driven by more effective teachers being more likely to obtain certification. Although certification does not seem to be indicative of teacher quality, principals are shown to be able to identify teachers at the top and the bottom of the teacher quality distribution (Jacob and Lefgren 2008) and that it may be easier to pick a good teacher than to train one (Chingos and Peterson 2011; Rockoff et al. 2011).

Another strand of research investigates how teachers perform in examinations themselves. Harris and Sass (2011) find no correlation between SAT performance and classroom effectiveness in Florida, while Boyd et al. (2008) find a positive correlation with mathematics SAT scores in New York City. Clotfelter et al. (2006) find a very weak correlation between licensure test performance and classroom effectiveness in North Carolina. However, several studies discover significant relationships between measures more closely attuned to the content knowledge used in teaching and student achievement (see the review in Baumert et al. (2010)). Hill, Rowan, and Ball (2005) find that teachers' mathematical knowledge is significantly related to student achievement

in the first and third grades in U.S. elementary schools. Metzler and Woessmann (2012) estimate the causal effect of teacher subject knowledge on student achievement using within-teacher within-student variation, exploiting a unique Peruvian sixth-grade dataset that tested students and their teachers in two subjects. This allows circumventing omitted-variable and selection biases by using student and teacher fixed effects and observing teachers teaching both subjects in one-classroom-per-grade schools. The findings suggest that a one standard deviation in subject-specific teacher achievement increases student achievement by about 10 percent of a standard deviation.

For Germany, the COACTIV study, which is embedded in the longitudinal component of the German PISA 2003 study, involves two assessments of students and their mathematics teachers at the end of grades 9 and 10. It provides evidence on the association between student achievement and different dimensions of teacher knowledge. Teachers were tested in mathematics-related content knowledge (conceptual and/or procedural mathematical skills) and mathematics-related pedagogical content knowledge (teachers' knowledge of tasks, student cognitions, and instruction). Results from COACTIV indicate significant relationships between student achievement and two types of teachers' knowledge: content knowledge and pedagogical content knowledge (Baumert and Kunter 2011; Baumert et al. 2010; Kunter et al. 2007).

There is also research more directly focused on the actual teaching process and the impact of in-class activities on student achievement (for reviews of this literature, see Seidel and Shavelson (2007) and Slavin, Lake, and Groff (2009)). The findings of this literature underscore the importance of teaching practices, instructional skills, and classroom management for student performance. For example, Kane, Staiger, and Rockoff (2010) and Tyler et al. (2010) find that classroom management and instructional skills as measured by the Teacher Evaluation System in Cincinnati can predict student achievement. Their classroom observation measures capture teaching practices such as "the teacher establishes effective routines and procedures … and manages transitions to maximize instructional time" and "the teacher engages students in discourse and uses thought-provoking questions." Lavy (2011) finds that teaching emphasizing in-the-classroom instilment of knowledge and comprehension has a very strong and positive effect on test scores. A meta-analysis by Slavin et al. (2009) reveals significant impacts of cooperative learning programs in mathematics instruction that target teachers' instructional behaviors rather than mathematics content alone. Lou et al. (1996) argue that within-class grouping, a typical component of cooperative learning strategies, has potential to enhance student achievement.

However, another meta-analysis by Dignath and Buettner (2008) reveals negative impacts of group work for primary school students. Based on data from the National Educational Longitudinal Study, Goldhaber and Brewer (1997) find that instruction in small groups and emphasis on problem solving is associated with lower student mathematics test scores for 10th-grade students. Aslam and Kingdon (2011) analyze student achievement data for Pakistan and find that students have higher test scores when taught by teachers who spend more time on lesson planning and ask more questions in class, while Schwerdt and Wuppermann, (2011) show that traditional teaching practices may be beneficial for students in contrast to general beliefs. In his quasi-experimental evidence on class-size effects in Europe, Woessmann (2005) shows that the impact of class size on student achievement decreases with teacher quality.

2.2 Educational Institutions

In Germany, teaching is organized in classes, rather than by courses and, in general, all students in one classroom receive teaching from the same teacher in a given subject, thus not affecting estimation due to within-classroom tracking. Furthermore, teachers do not specialize in teaching one specific grade, but are assigned by school management to certain classes on a yearly basis. Teachers teaching a certain subject in Grade 5 are therefore quite likely to teach the same class in Grade 6.

2.2.1 Prospective Teachers' Transition from High School to University

Secondary education is tracked in Germany and only specific tracks give access to tertiary education. The degrees that give access to universities are all designed quite similarly and are earned during the last two years of high school and by the final examinations. Students are somewhat free to choose a set of courses and their duration and receive a final grade every semester for every class they have taken. Students additionally choose to be tested in four or five subjects by somewhat standardized final examinations. A weighted average of grades earned each semester, the grade received on one term paper, and final examinations form the high school GPA.

Conditional on having earned one degree giving access to university, enrollment in tertiary education is either open or almost exclusively based on the high school GPA.[1] After a student who

[1] Some programs also give weight to motivational letters or extracurricular activities.

wishes to become a teacher receives her high school degree, she applies to university programs that are designed to determine the school track, the subjects, and the state in which she will eventually teach. This is different from the U.S. system in two important ways. First, in Germany, students do not have a first year at university during which they receive general education in various fields. Second, to a large extent the German teaching program in which a student enrolls determines the future employment trajectory of a prospective teacher.

As the high school GPA includes grades received during two years of high school and those received on four or five final examinations, it is quite rich in information. It is also quite different from the U.S. SAT in that the grades included have been received for written and oral examinations, presentations, term papers, and final examinations in a broad range of subjects and thus the German high school GPA can be regarded as a measure of general education.

2.2.2 Teacher Training for Secondary Education

In Germany, 75 to 80 percent of teachers are graduates of a formal teacher education program (BMBF 2012). High school graduates who decide to enter a secondary education teaching program may occasionally have to fulfill certain entry requirements for a university, dependent on their high school GPA, and must choose a program that is specific to a state, a school type, and (at least) two academic subjects.

Once entered, teacher training for secondary education in Germany takes place in two steps, with the structure and content of training varying partially on the state level. Generally, the first phase takes place at a university and lasts four to six years, depending on the state. The courses include (at least) two subjects that will later be taught, pedagogics, and internships at schools. At the end of the first phase, student teachers must take exams that measure theoretical knowledge in the taught subjects and pedagogics. The outcome of these exams and the grades earned at the university level (weighted by class credits) comprise the first state examination (FSE) grade.

The second stage of teacher training involves a one-and-a-half to two-year practical program of teacher seminars at teacher training schools. During this phase, every student teacher is given a teaching position. Trainee teachers are employed and teach regular classes. During this phase, trainee teachers must complete a thesis, pass several oral examinations in the subjects taught, and present three demonstration lessons that are rated by head teachers. The second state examination

(SSE) grade is based on the thesis grade, the oral exams, and the assessments of the demonstration lessons.

2.2.3 Teacher Labor Market

Entry into the profession is based on the supply of teachers and the demand for them by schools. Generally, teaching degrees specify the subjects to be taught, the type of school, and the state, and the markets are divided accordingly. For each cohort of student teacher graduates, each market clears on the basis of the (weighted) grade point average in the two state examinations.

Teachers who successfully enter the profession rarely exit before retirement. The mean leaving age for men is 60, for women 55, with medians of 62 and 60, respectively (BMBF 2012, Table 3.2). The two most common reasons besides retirement for leaving (temporarily) are the birth of a child and long-term illness, which explains why the leaving age distribution is skewed. Most teachers are civil servants, and teacher pay is regulated at the state level, based largely on tenure and partially on assessments by principals.

3. Data: The German National Education Panel Study

I use data from the German National Educational Panel Study (NEPS). NEPS is an education research project in Germany that began in 2010. The project draws from a representative sampling of individuals from six starting cohorts: newborns, kindergarten students, 5th grade students, 9th grade students, university students and adults. The data used in this study stems from the starting cohort of 5th grade students.[2] It includes, besides students' testing and background data, information from parents' and principals' questionnaires. Additionally, and most importantly, it includes extensive information about mathematics and language teachers for the given school year.[3]

Figure 1 provides a graphical overview of the relevant data for this study. I use data gathered from the first three waves of this cohort. Students in that cohort were sampled using a stratified sampling procedure. Schools were randomly drawn from the population of public schools to be

[2] This cohort is also referred to as starting cohort 3 (SC3).

[3] For a detailed, general overview of the NEPS see Blossfeld, Roßbach, and von Maurice (2011) and for the competence tests in particular see NEPS (2011a, 2011b).

representative by school type. From the selected schools, two classrooms (if available) were randomly asked to participate in the study. For students, participation in the study involves testing and completing a questionnaire. Questionnaires for relevant context persons (parents, homeroom teachers, principals and teachers in mathematics and German) were also required. For each student and his or her context persons, participation is voluntary.

In addition to the testing information, data from student and teacher questionnaires was also used for the main estimation specifications. The student questionnaires give insight into socio-economic background and are less affected by attrition than the parental data, making them the optimal choice in the tradeoff between covariate availability and representativeness.[4] Table 1 shows a selection of student background information by school track that is later used as covariates and also includes the outcome variables that are standardized to have a mean of zero and a standard deviation of one. 49% of students are female with an average age of 10.76. About 8% of students have stayed down a year at least once.

The teacher questionnaires, in contrast to administrative data from the United States, are extensive and provide information about teachers' demographics, philosophies, educational goals, stress in the profession, colleagues, perception of the profession, participation in extracurricular activities and further training, aspects of career choice, certification, study history, subjects taught, high school GPA, and state examination grades. Table 2 provides information about basic teacher characteristics by subject taught. For a subset of teachers, data about in-class time use is also available and will be analyzed in Section 6. One can see that selection into the teaching profession in Germany shows a tendency towards women with about 74% in the full sample with a higher share of men teaching math relative to German. Teachers with a migration background (at least one parent born abroad) are rather rare with 5 percent. Two thirds of teachers start their teacher training right after graduating from high school and three quarters possess previous teaching experience (e.g. private lessons). One in five teachers at least started another field of study at some point. The grades as the main variables of interest are linearly transformed to be better interpretable:[5] A higher score meaning a better grade, ranging from 4 (A) to 0 (F).

[4] Besides item non-response, about half of the sample cannot be matched to data from both of the teachers (see Figure A1). Investigating the pattern of attrition, I solely find evidence for a slight positive selection of the sample as measured by the previous math grade and migration background.

[5] German grades generally range discretely from 1 (very good) to 6 (insufficient).

The outcome variables are scores in standardized tests in mathematics and language, standardized to have a mean of zero and a standard deviation of one. Testing in the first wave took place in November and December of 2010, near the beginning of the school year. For my main specifications, I use data from students that were tested in the same skill domains in the beginning of grade 5 and grade 7.

4. The Value-Added Model

The starting point of my main analyses is a slight variation of the educational production framework developed in Todd and Wolpin (2003), which describes the process of cognitive skill production as follows:

$$y_{it}^{k} = y_{t}^{k}[\Gamma_{TS,j}, X_{i}^{k}, \mu_{i}^{k}, \varepsilon_{it}^{k}] \quad (1)$$

Cognitive achievement of individual i at time t in the cognitive dimension of subject k is a function defined by the technology $y_t^k(.)$ that translates the educational inputs of the vector of the entire history of all family and school (X) and teacher (Γ_{TS}) inputs until time t, as well as the subject specific time-invariant mental endowment μ_i^k into the educational outcome.

To investigate the relationship between teacher effectiveness and teacher skills, I replace the teacher FE component with the respective trait T as in:

$$y_{i7}^{k} = \lambda^{k} y_{it}^{k} + \beta^{k} \times T_{j(5,6)}^{k} + \gamma X_{i}^{k} + \varepsilon_{i7}^{k} \quad (2)$$

The test score y of student i at the beginning of grade 7 in the skill domain $k \in (math; language)$ is a function of its first or second lag $t \in (1; 2)$, teacher j's trait T, and i's individual and family inputs. Note that the sample is limited to students who were taught by the same teacher in both Grades 5 and 6, as otherwise the teacher components cannot be distinguished (Leigh 2010).

The direct estimate of the effect of teacher characteristics on student achievement produced by Equation (2) is not biased by between- or within-school sorting of students based on unobservable student traits. However, to interpret the estimated vector of coefficients β causally, one would have to make the identifying assumption that unobservable teacher characteristics that

directly influence student achievement are not related to the observed teacher characteristics. This is a strong assumption and one that is likely to be violated. Hence, I interpret the estimates as correlates of teacher quality, which is standard in the literature on teacher effectiveness.

One potential threat to identification may be that students on high-achievement trajectories are assigned to teachers based on observable teacher traits. However, this is unlikely, as almost all students in the sample had just switched from primary to secondary school and their parents therefore do not yet have enough knowledge about the new school environment to successfully manipulate teacher or classroom assignment. Additionally, Fischer and Enzi (2016) provide suggestive evidence that, conditional on attending a particular school, classroom assignment is random.

Nor are the results confounded by selection into schools or by time-invariant student subject-specific ability component, as the standard VA model accounts for these aspects by including the lagged test score that captures both, as in a standard first difference specification.

5. Main Results

All standard errors and test statistics are based on cluster-robust standard errors at the school level—the initial stage of the stratified sampling procedure of NEPS.[6] Table 3 and Table 4 show the main regression results of math and language teachers, respectively, and their high school GPA and state examinations for the two-year gain specification. Table 5[7] shows the same results for language teachers, but for the skill domain of vocabulary in Grade 6, holding constant reading speed, comprehension, and orthography in Grade 5. If these three test scores from Grade 5 jointly cover the hypothetical Grade 5 vocabulary skill, all the exclusion restrictions of a regular VA model hold.

Columns (1), (4) and (7) are simple OLS regressions including solely one teacher grade at a time and the lagged test score from the beginning of Grade 5. Columns (2), (5) and (8) add student covariates, and Columns (3), (6) and (9) add Grade 5 test scores in perception speed, logic, and either math or language. Finally, Column (10) puts all teacher grades in one specification. Due to

[6] Clustering at the classroom level does not substantially alter the standard errors.

[7] Table A1 shows the same results for another cohort from NEPS. For a teacher's high school GPA, the same results emerge. For the second state examination, the estimates point towards the same direction, however statistically insignificantly.

a comparably small sample, item non-response, and somewhat high correlations between grades (ranging from 0.25 to 0.5) my preferred specifications are those of Columns (3), (6) and (9). All teacher grades are standardized to have a mean of zero and a standard deviation of one, as are the student test scores. The estimates can be interpreted as follows: a one SD increase in a teacher's grade changes student gains by the estimate measured in SD.

High school GPA is an economically and statistically significant determinant of teacher effectiveness in almost all specifications and in both subjects. Table 3 starts out with an estimate of 0.06, decreasing to 0.04 when adding basic student covariates, which is not statistically different from the original estimate. Adding covariates without the estimate substantially changing supports the assumption of conditional random assignment of teachers to classrooms. Adding further test results yields my preferred estimate of 0.035; thus a teacher having a one SD higher high school GPA is associated with an increase in student achievement of 3.5% SD.

The exact same pattern can be found in Table 4 for two-year gains in students' language achievement, albeit with less explanatory power, as is usual in research on determinants of teacher effectiveness as it is harder to measure language skills or harder to influence them. Evidence for the first scenario can be found in the substantially lower R^2 throughout all specifications compared to the math results. Less persistence in language skills is another, but unlikely, explanation.

The pattern of the high school GPA being a strong determinant reappears for one-year language gains, which is most likely due to the fact that I gain statistical power due to the larger number of observations that is gained by not conditioning the sample on students that were taught by the same teacher in two consecutive school years. This finding is probably not due to less measurement error or higher persistence in this skill domain either because neither the coefficient of determination nor the standard errors of the marginal effects estimates are considerably different from the previous one. Even with conditioning on a teacher's state exam grades, the high school GPA preserves its strong predictive power for student achievement in Table 5.

With the exception of Column (4) of Table 5, the grade of the first state examination (FSE) does not yield any statistically significant results; however, the size of the estimates is generally not different from the ones of the second state examination (SSE). The estimates range from 1.8% to 3.6% of a SD in mathematics and from -2.4% to 5.6% of a SD in German, which likely stems from the fact of larger noise in the measurement of language skills compared to mathematical competencies. In columns (6) of Tables 4 and 5 the preferred estimates of language teacher's FSE

are shown: The effect ranges from -2.4% to 2.6% with standard errors of 0.02 yielding a span of potential parameters somewhere in between about -6% to 6% for an increase of a SD in a teacher's FSE.

The preferred estimate of mathematics teachers' SSE grade can be found in Table 3's column (9). A one SD increase in SSE increases student achievement by 2.5% of a SD, with a rather small standard error implying a 95% confidence interval of 0% to 5% of a SD. For language teachers, the results look mixed in Tables 4 and 5: On the one hand, the results vary from -5.1% to -1.2% in Table 4 for two-year gains, on the other hand, from 0% to 7% for one-year gains in Table 5, implying a broad range of potential true parameters. The preferred results of columns (9) show estimates from -3% to 3% with standard errors of 0.02.

Summing up, I find evidence for a positive effect of teachers with a higher high school GPA in both subjects and a positive effect for mathematics teachers with a higher SSE grade.

6. Pre-Service Exams and In-Class Time Use

The results found for good grades in Abitur and on the SSE may be because these teachers act substantially different in the classroom than their lower-scoring counterparts. Hence, correlations of pre-service exam results and in-class time use may shed some light on the underlying forces at work. Simple regression results are presented in Table 6 for a subset of math and language teachers who were, in addition to the basic questionnaire, asked about the share of time they commit to certain activities within the classroom. Column (1) shows the result for the share of time spent discussing homework, Column (2) for teacher presentations, and Columns (3) and (4) for tasks with and without assistance, respectively. Column (5) represents time spent on repetitive drills, Column (6) on taking tests, Column (7) on classroom management, and Column (8) on other activities.

There are mixed results, including that better teachers, in terms of Abitur and SSE grades, need less time for classroom management. Hence, more effective teachers seem to be so because they need less time for classroom management and hence have more time for other activities, which are, in turn, classroom specific and not of the "one size fits all" variety and hence in contrast to the findings of Schwerdt, and Wuppermann (2011) that find traditional teaching to be beneficial for students.

7. Heterogeneous Effects and Nonlinearities

Some subgroups of students have been shown to especially benefit from certain teacher traits (Dee 2005), and girls have been shown to benefit from a single-sex classroom in mathematics. Hence, it may be that the main results are driven by some teachers adjusting better to girls' needs in mathematics class. Table 7 and Table 8 show results for boys versus girls and pupils with a migration background versus natives, respectively.

Table 7 makes it clear that girls gain less in mathematics over time than do boys (see line labeled "Female"). If anything, it seems that girls benefit more from teachers with higher grades in high school and on the FSE, whereas boys benefit more from teachers with higher teaching capabilities as measured by the SSE.

Table 8 reveals that students with a migration background tend to gain less over time. Adding interaction terms of migration status and teachers' grades yields no statistically significant findings for migrants, although there is a tendency toward a negative correlation. However, inclusion of the migration status interaction reveals positive and significant findings for the native population for Abitur, FSE, and SSE, with SSE having the most consistent such effect. Natives are more affected by the SSE than are migrants: A one SD increase in a teacher's grade in SSE leads to 3.2% higher annual gains in math test scores for them.

I take a deeper look at the discovered effects of the state examinations and teachers' high school GPA by dividing these variables into quartiles and searching for potential nonlinearities, which are likely according to Jacob and Lefgren (2008), who found that principals can more easily tell who are the best and worst teachers, but are not doing so well at telling mediocre teachers apart from each other. Table 9 shows that the results in mathematics stem from the top teachers: being in the top quartile of the high school GPA is linked with a 12.1% increase in student achievement, for the FSE it is 7.7% and for the SSE it is 7.3%. Hence all three grades are good at distinguishing the best from the rest when it comes to mathematics teachers. No such pattern exists for language teachers.

8. Conclusion

This study investigates the effects of pre-service cognitive and pedagogical skills on teacher effectiveness. German teacher training and selection provides an interesting case study to investigate the impact of teacher's academic and practical education as well as their selection: In

a first stage, high school graduates enroll in a university program that is sometimes defined by a specific high school GPA cutoff so that below a certain threshold high school graduates cannot enroll in some programs. The choice of such a program is already very consequential: It defines the state, the subjects and the school track that the potential teacher is going to work in later on. Once having entered the university program, a regular academic training ensues up until the first state examination that measures the cognitive skills of the student teachers. Having passed that exam, the student teacher can go on to become a trainee teacher that is teaching real classrooms under the supervision of a head teacher. That head teacher – with a committee of other head teachers – grades the trainee teacher after one and a half to two years of practical experience, largely based on two to three demonstration lessons. Hence these three grades, on the one hand, define almost exclusively the teacher selection of German states and, on the other hand, they provide information about the teacher's academic and practical skills.

To formalize a teacher's quality, I use a standard value-added model of student achievement to capture the part of the variation in a student's test score gain that is solely based on a teacher's effectiveness and correlate it with the above mentioned grades. I use rich micro-level data from the German National Educational Panel Study (NEPS) that comprises a nationally representative sample of secondary school students.

I find that being in the top quartile of the three measures of cognitive and pedagogical skills is linked with higher teacher effectiveness in mathematics, while it is not for language teachers. A higher high school GPA is linked with needing less time for classroom management in mathematics. Subgroups of students have been shown to benefit differentially from teacher traits, however no such pattern exists in dividing students by gender and migration status.

References

Aslam, M., and G. Kingdon. 2011. "What Can Teachers Do to Raise Pupil Achievement?" *Economics of Education Review* 30(3):559–74.

Baumert, J. et al. 2010. "Teachers' Mathematical Knowledge, Cognitive Activation in the Classroom, and Student Progress." *American Educational Research Journal* 47(1):133–80.

Baumert, J., and M. Kunter. 2011. "Das mathematikspezifische Wissen von Lehrkräften, kognitive Aktivierung im Unterricht und Lernfortschritte von Schülerinnen und Schülern." Pp. 163–92 in *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV*, edited by M Kunter et al. Münster: Waxmann.

Blossfeld, H.-P., H.-G. Roßbach, and J. von Maurice. 2011. Education as a Lifelong Process: The German National Educational Panel Study (NEPS). *Zeitschrift für Erziehungswissenschaft*: Special Issue 14.

BMBF. 2012. *Systematisierung der Lehrerforschung und Verbesserung ihrer Datenbasis*. Berlin.

Boyd, D., H. Lankford, S. Loeb, J. Rockoff, and J. Wyckoff. 2008. "The Narrowing Gap in New York City Teacher Qualifications and Its Implications for Student Achievement in High-Poverty Schools." *Journal of Policy Analysis and Management* 27(4):793–818.

Cantrell, S., J. Fullerton, T. J. Kane, and D. O. Staiger. 2008. *National Board Certification and Teacher Effectiveness: Evidence from a Random Assignment Experiment*. NBER Working Paper 14608. Cambridge, MA.

Chetty, R., J.N. Friedman, and J.E. Rockoff. 2014. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review* 104(9):2593–632.

Chingos, M.M., and P.E. Peterson. 2011. "It's Easier to Pick a Good Teacher than to Train One: Familiar and New Results on the Correlates of Teacher Effectiveness." *Economics of Education Review* 30(3):449–65.

Clotfelter, C. T., H. F. Ladd, and J. L. Vigdor. 2006. "Teacher-Student Matching and the Assessment of Teacher Effectiveness." *Journal of Human Resources* 41(4):778–820.

Dee, T.S. 2005. "A Teacher like Me: Does Race, Ethnicity, or Gender Matter?" *American Economic Review* 95(2): 58–165.

Dignath, C., and G. Buettner. 2008. “Components of Fostering Self-Regulated Learning among Students: A Meta-Analysis on Intervention Studies at Primary and Secondary School Level.” *Metacognition and Learning* 3:231–64.

Fischer, M., and B. Enzi. 2016. *Classroom Composition and Academic Self-Concept in Educational Production*. Mimeo.

Goldhaber, D. 2002. “The Mystery of Good Teaching.” *Education Next* 2(1):50–55.

Goldhaber, D., and E. Anthony. 2007. “Can Teacher Quality Be Effectively Assessed? National Board Certification as a Signal of Effective Teaching.” *Review of Economics and Statistics* 89(1):134–50.

Goldhaber, D., and D.J. Brewer. 1997. “Why Don’t Schools and Teachers Seem to Matter? Assessing the Impact of Unobservables on Educational Productivity.” *Journal of Human Resources* 505–23.

Hanushek, E. A., and S. G. Rivkin. 2006. “Teacher Quality.” Pp. 1051–78 in *Handbook of the Economics of Education*, edited by E A Hanushek and F Welch. Amsterdam: North Holland.

Harris, D. N., and T. R. Sass. 2011. “Teacher Training, Teacher Quality and Student Achievement.” *Journal of Public Economics* 95(7-8):798–812.

Hill, H.C., B. Rowan, and D. Loewenberg Ball. 2005. “Effects of Teachers’ Mathematical Knowledge for Teaching on Student Achievement.” *American Educational Research Journal* 42(2):371–406.

Jacob, B. A., and L. Lefgren. 2008. “Can Principals Identify Effective Teachers ? Evidence on Subjective Performance Evaluation in Education.” *Journal of Labor Economics* 26(1): 101–36.

Kane, T. J., D. O. Staiger, and J. E. Rockoff. 2010. “Searching for Effective Teachers with Imperfect Information.” *Journal of Economic Perspectives* 24(3):97–118.

Kunter, M. et al. 2007. “Linking Aspects of Teacher Competence to Their Instruction: Results from the COACTIV Project.” Pp. 39–59 in *Studies on the Educational Quality of Schools: The Final Report on the DFG Priority Programme*. Waxmann.

Ladd, H. F., T. R. Sass, and D. N. Harris. 2007. *The Impact of National Board Certified Teachers on Student Achievement in Florida and North Carolina*.

Lavy, V. 2011. *What Makes an Effective Teacher - Quasi-Experimental Evidence*. NBER Working Paper 16885. Cambridge, MA.

Leigh, A. 2010. "Estimating Teacher Effectiveness from Two-Year Changes in Students' Test Scores." *Economics of Education Review* 29(3):480–88.

Lou, Y. et al. 1996. "Within-Class Grouping: A Meta-Analysis." *Review of Educational Research* 66(4):423–58.

Metzler, J., and L. Woessmann. 2012. "The Impact of Teacher Subject Knowledge on Student Achievement: Evidence from within-Teacher within-Student Variation." *Journal of Development Economics* 99(2):486–96.

NEPS. 2011a. *Starting Cohort 3 - Information on the Competence Test*. Bamberg.

NEPS. 2011b. *Starting Cohort 4 Students - Information on the Competence Test*. Bamberg.

Rivkin, S. G., E. A. Hanushek, and J. F. Kain. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica* 73(2):417–58.

Rockoff, J.E., B.A. Jacob, T.J. Kane, and D.O. Staiger. 2011. "Can You Recognize an Effective Teacher When You Recruit One?" *Education Finance and Policy* 6(1):43–74.

Schwerdt, G., and A.C. Wuppermann. 2011. "Is Traditional Teaching Really All That Bad? A within-Student between-Subject Approach." *Economics of Education Review* 30(2):365–79.

Seidel, T., and R. J. Shavelson. 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):454–99.

Slavin, R. E., C. Lake, and C. Groff. 2009. "Effective Programs in Middle and High School Mathematics: A Best-Evidence Synthesis." *Review of Educational Research* 79(2):839–911.

Todd, P.E., and K.I. Wolpin. 2003. "On the Specification and Estimation of the Production Function for Cognitive Achievement." *Economic Journal* 113(485):F3–F33.

Tyler, J.H., E.S. Taylor, T.J. Kane, and A.L. Wooten. 2010. "Using Student Performance Data to Identify Effective Classroom Practices." *American Economic Review* 100(2):256–60.

Woessmann, L. 2005. "Educational Production in Europe." *Economic Policy* 20(43):446–504.

Figures and Tables

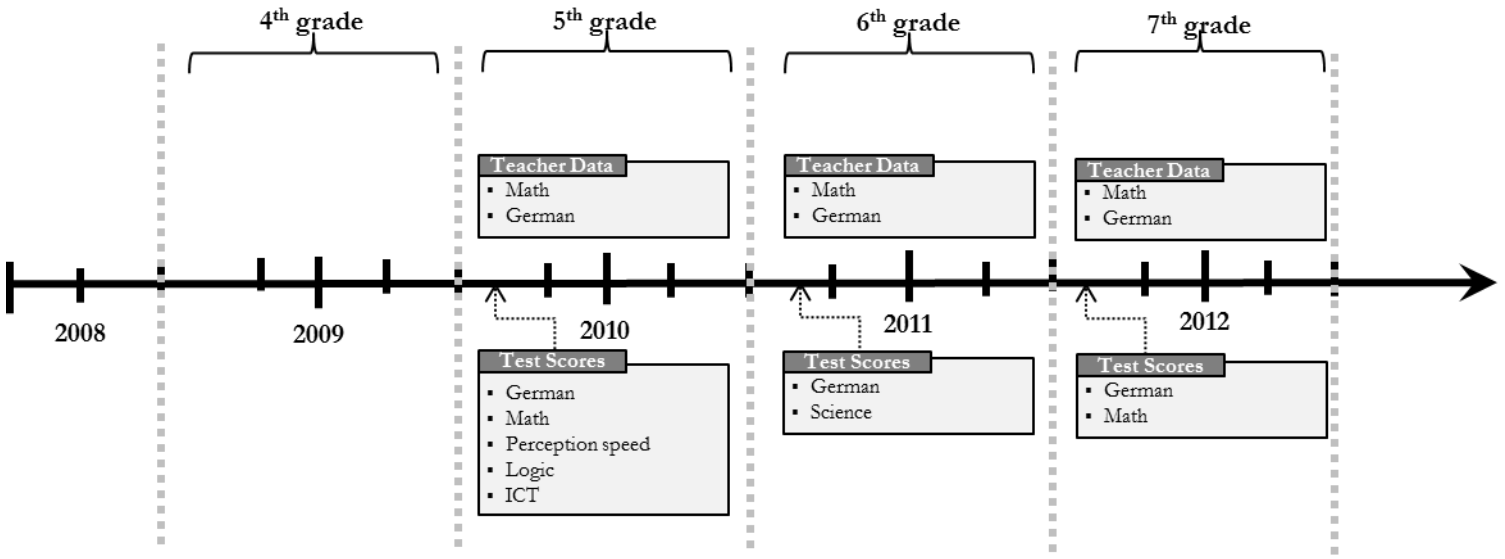

Figure 1: Timing of testing and available teacher data of 5th grade students

Table 1: Descriptive statistics of 5th grade students

Female	0.49	(0.50)
Age	10.76	(0.51)
Born abroad	0.04	(0.20)
Household size	4.65	(1.73)
Mother born abroad	0.19	(0.39)
Father born abroad	0.19	(0.39)
PC availability	1.70	(0.52)
Retention	0.08	(0.28)
Number of books	3.77	(1.38)
Perceptual speed score	43.45	(13.05)
Reasoning score	6.94	(2.57)
German: Grade Point	2.32	(0.88)
Math: Grade Point	2.27	(0.94)
German: Test score	0.00	(1.00)
Math: Test score	0.00	(1.00)
Observations	2234	

Unit of observation: Students. Means of variables with standard deviations in parentheses for the full sample of 5th grade students.

Table 2: Basic descriptive statistics of teachers by subject taught

	Math	German	Full
Female	0.69	0.79	0.74
	(0.46)	(0.40)	(0.44)
Born before 1950	0.05	0.04	0.05
	(0.22)	(0.21)	(0.21)
Born in 1950's	0.33	0.31	0.32
	(0.47)	(0.46)	(0.47)
Born in 1960's	0.23	0.22	0.22
	(0.42)	(0.41)	(0.42)
Born in 1970's	0.24	0.28	0.26
	(0.43)	(0.45)	(0.44)
Born after 1979	0.16	0.15	0.15
	(0.37)	(0.35)	(0.36)
Migration background	0.06	0.05	0.05
	(0.24)	(0.21)	(0.22)
East German	0.20	0.19	0.19
	(0.40)	(0.39)	(0.40)
Imm. start studies	0.66	0.64	0.65
	(0.47)	(0.48)	(0.48)
Previous teaching experience	0.74	0.69	0.72
	(0.44)	(0.46)	(0.45)
Other field of study	0.20	0.20	0.20
	(0.40)	(0.40)	(0.40)
GPA: High School	2.75	2.64	2.70
	(0.63)	(0.58)	(0.61)
Grade: 1st State exam	2.87	2.96	2.91
	(0.53)	(0.55)	(0.54)
Grade: 2nd State exam	2.94	2.99	2.96
	(0.59)	(0.58)	(0.59)
Observations			740

Unit of observation: Teachers. Means of variables with standard deviations in parentheses.

Table 3: Main regression results of math teachers' pre-service cognitive and pedagogical skills

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
GPA: High School	0.063***	0.044**	0.035*							0.026
	(0.023)	(0.021)	(0.019)							(0.023)
Grade: 1st State Exam				0.036	0.022	0.018				-0.011
				(0.022)	(0.020)	(0.017)				(0.026)
Grade: 2nd State Exam							0.036**	0.023	0.025*	0.018
							(0.018)	(0.015)	(0.013)	(0.024)
L. Student TS Math	0.733***	0.739***	0.730***	0.573***	0.564***	0.558***	0.398***	0.387***	0.384***	0.378***
	(0.020)	(0.021)	(0.022)	(0.021)	(0.021)	(0.022)	(0.024)	(0.024)	(0.026)	(0.029)
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes
Other tests	No	No	Yes	No	No	Yes	No	No	Yes	Yes
Observations	2329	2329	2328	2383	2383	2381	2128	2128	2127	1742
R^2	0.535	0.582	0.622	0.536	0.586	0.625	0.520	0.572	0.610	0.609

Dependent variable: Math test score in grade 7. Standard errors in parentheses. Controls include a student's gender, year, month and place of birth, parents' place of birth, PC availability, grades in math and language, peers' attitude toward education, school and classroom share of migrants, household size, grade retention status, number of books at home, religious denomination and SDQ[8] scores. Other tests include grade 5 tests in perception speed, logic, ICT and math or German test scores respectively. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

[8] The Strengths and Difficulties Questionnaire (SDQ) is a brief behavioral screening questionnaire about 3-16 year olds.

Table 4: Main regression results of language teachers' pre-service cognitive and pedagogical skills

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
GPA: High School	0.049*	0.013	-0.016							0.004
	(0.026)	(0.025)	(0.023)							(0.034)
Grade: 1st State Exam				0.001	-0.016	-0.024				0.005
				(0.030)	(0.026)	(0.023)				(0.036)
Grade: 2nd State Exam							-0.012	-0.035	-0.033	-0.051*
							(0.028)	(0.024)	(0.020)	(0.026)
L. Student TS R.Comp	0.630***	0.522***	0.319***	0.637***	0.518***	0.309***	0.636***	0.508***	0.316***	0.331***
	(0.022)	(0.021)	(0.023)	(0.021)	(0.021)	(0.024)	(0.024)	(0.024)	(0.024)	(0.025)
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes
Other tests	No	No	Yes	No	No	Yes	No	No	Yes	Yes
Observations	1939	1939	1937	2083	2083	2079	1701	1701	1698	1428
R^2	0.385	0.426	0.497	0.390	0.434	0.509	0.384	0.432	0.500	0.495

Dependent variable: German test score in grade 7. Standard errors in parentheses. Controls include a student's gender, year, month and place of birth, parents' place of birth, PC availability, grades in math and language, peers' attitude toward education, school and classroom share of migrants, household size, grade retention status, number of books at home, religious denomination and SDQ scores. Other tests include grade 5 tests in perception speed, logic, ICT and math or German test scores respectively. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Main regression results of language teachers' pre-service cognitive and pedagogical skills for one-year gains

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
GPA: High School	0.133***	0.089***	0.069**							0.086***
	(0.035)	(0.030)	(0.029)							(0.033)
Grade: 1st state exam				0.056**	0.035	0.026				-0.006
				(0.027)	(0.021)	(0.020)				(0.027)
Grade: 2nd state exam							0.066***	0.028	0.026	0.006
							(0.023)	(0.022)	(0.021)	(0.026)
L. Student TS R. Speed	0.094***	0.078***	0.058***	0.107***	0.087***	0.073***	0.095***	0.079***	0.061***	0.071***
	(0.022)	(0.020)	(0.019)	(0.022)	(0.020)	(0.020)	(0.022)	(0.020)	(0.019)	(0.019)
L. Student TS Orthography	0.013	-0.004	-0.061***	0.006	-0.013	-0.072***	-0.003	-0.031	-0.080***	-0.065**
	(0.021)	(0.021)	(0.022)	(0.023)	(0.022)	(0.023)	(0.022)	(0.023)	(0.022)	(0.025)
L. Student TS R. Compreh.	0.499***	0.411***	0.280***	0.501***	0.403***	0.266***	0.511***	0.408***	0.279***	0.288***
	(0.026)	(0.022)	(0.022)	(0.025)	(0.020)	(0.021)	(0.027)	(0.023)	(0.024)	(0.027)
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes
Other tests	No	No	Yes	No	No	Yes	No	No	Yes	Yes
Observations	2285	2285	2282	2426	2426	2421	1963	1963	1959	1619
R^2	0.329	0.454	0.492	0.309	0.451	0.491	0.317	0.450	0.488	0.494

Dependent variable: German test score in Grade 6. Standard errors in parentheses. Controls include a student's gender, year, month, and place of birth, parents' place of birth, PC availability, grades in math and language, peers' attitude toward education, school and classroom share of migrants, household size, grade retention status, number of books at home, religious denomination, and SDQ scores. Other tests include Grade 5 tests in perception speed, logic, ICT, and math or German test scores, respectively.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: OLS regression results for the relationship of in-class time use and teachers' skills

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Math teacher sample (N=232)								
GPA: High School	-0.372	0.175	0.400	-1.250	1.152	0.843*	-1.327***	-0.331
	(0.608)	(0.769)	(1.014)	(1.104)	(0.803)	(0.434)	(0.417)	(0.627)
Grade: 1st state exam	0.337	-1.719**	0.934	0.900	0.182	-0.194	0.333	-0.483
	(0.560)	(0.848)	(0.952)	(1.064)	(0.836)	(0.420)	(0.428)	(0.551)
Grade: 2nd state exam	0.602	-0.200	-0.696	-0.595	0.078	0.154	0.005	0.448
	(0.558)	(0.704)	(0.986)	(1.180)	(0.814)	(0.429)	(0.413)	(0.565)
German teacher sample (N=384)								
GPA: High School	-0.471	0.766	0.042	-0.383	-0.316	0.382	-0.647	0.564
	(0.564)	(0.579)	(0.715)	(1.026)	(0.695)	(0.436)	(0.398)	(0.434)
Grade: 1st state exam	-0.011	-0.351	0.370	0.921	-0.201	0.528	-0.451	-0.575
	(0.645)	(0.654)	(0.655)	(1.063)	(0.654)	(0.410)	(0.407)	(0.465)
Grade: 2nd state exam	-0.205	-0.773	-0.281	1.221	0.320	0.584	-0.735**	-0.766*
	(0.617)	(0.637)	(0.582)	(0.889)	(0.645)	(0.369)	(0.369)	(0.447)

Dependent variables: one dimension of in-class time use as a share of overall time available for teaching. Standard errors in parentheses. Each cell represents a distinct regression result. Column (1) covers time spent discussing homework, Column (2) teacher presentations, Columns (3) and (4) tasks with and without assistance, respectively. Column (5) represents repetitive drills, Column (6) taking tests, Column (7) classroom management, and Column (8) other activities. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Heterogeneous effect of math teachers' pre-service cognitive and pedagogical skills by students' gender

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Female	-0.121***	-0.122***	-0.140***	-0.143***	-0.150***	-0.173***	-0.177***	-0.176***	-0.206***	-0.176***
	(0.029)	(0.028)	(0.029)	(0.029)	(0.026)	(0.027)	(0.029)	(0.026)	(0.027)	(0.032)
GPA: High School	0.053**	0.031	0.022							0.031
	(0.026)	(0.022)	(0.019)							(0.025)
Grade: 1st State Exam				0.030	0.009	0.002				-0.048
				(0.031)	(0.027)	(0.024)				(0.032)
Grade: 2nd State Exam							0.051**	0.029	0.030	0.047
							(0.024)	(0.021)	(0.019)	(0.030)
F x GPA: High School	0.024	0.026	0.027							-0.007
	(0.027)	(0.025)	(0.026)							(0.031)
F x 1st StE				0.019	0.026	0.031				0.076**
				(0.032)	(0.030)	(0.028)				(0.038)
F x 2nd StE							-0.025	-0.012	-0.010	-0.059
							(0.028)	(0.027)	(0.026)	(0.038)
L. TS	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes
Other tests	No	No	Yes	No	No	Yes	No	No	Yes	Yes
N	2329	2329	2328	2383	2383	2381	2128	2128	2127	1742
R2	0.54	0.58	0.62	0.54	0.59	0.63	0.52	0.57	0.61	0.61
Full effect	0.077	0.057	0.048	0.048	0.035	0.033	0.026	0.016	0.020	
p	0.008	0.042	0.062	0.037	0.124	0.095	0.208	0.409	0.260	

Dependent variable: Math test score in grade 7. Standard errors in parentheses. Controls include a student's gender, year, month and place of birth, parents' place of birth, PC availability, grades in math and language, peers' attitude toward education, school and classroom share of migrants, household size, grade retention status, number of books at home, religious denomination and SDQ scores. Other tests include grade 5 tests in perception speed, logic, ICT and math or German test scores respectively. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Heterogeneous effect of math teachers' pre-service cognitive and pedagogical skills by students' migration background

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Migration background	-0.117***	-0.044	-0.037	-0.108***	-0.030	-0.029	-0.083*	-0.014	-0.017	-0.046
	(0.045)	(0.036)	(0.033)	(0.041)	(0.034)	(0.031)	(0.043)	(0.036)	(0.033)	(0.037)
GPA: High School	0.059**	0.036	0.029							0.007
	(0.026)	(0.023)	(0.020)							(0.025)
Grade: 1st State Exam				0.042**	0.028	0.024				-0.006
				(0.021)	(0.017)	(0.015)				(0.023)
Grade: 2nd State Exam							0.040**	0.029*	0.032**	0.028
							(0.019)	(0.016)	(0.015)	(0.025)
MB x GPA: High School	0.023	0.023	0.014							0.074*
	(0.040)	(0.035)	(0.031)							(0.037)
MB x 1st StE				-0.023	-0.031	-0.032				-0.029
				(0.044)	(0.039)	(0.033)				(0.050)
MB x 2nd StE							-0.018	-0.034	-0.034	-0.047
							(0.039)	(0.031)	(0.029)	(0.044)
L. TS	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes
Other tests	No	No	Yes	No	No	Yes	No	No	Yes	Yes
N	2313	2313	2312	2368	2368	2366	2118	2118	2117	1732
R2	0.54	0.58	0.62	0.54	0.59	0.62	0.52	0.57	0.61	0.61
Full effect	0.082	0.059	0.044	0.019	-0.003	-0.009	0.022	-0.005	-0.003	
p	0.036	0.084	0.133	0.695	0.943	0.804	0.538	0.868	0.915	

Dependent variable: Math test score in grade 7. Standard errors in parentheses. Controls include a student's gender, year, month and place of birth, parents' place of birth, PC availability, grades in math and language, peers' attitude toward education, school and classroom share of migrants, household size, grade retention status, number of books at home, religious denomination and SDQ scores. Other tests include grade 5 tests in perception speed, logic, ICT and math or German test scores respectively. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Nonlinearities in teachers' pre-service cognitive and pedagogical skills

	Math			Language		
	(1)	(2)	(3)	(1)	(2)	(3)
GPA: High School - 2nd Qt	0.081			0.025		
	(0.060)			(0.061)		
GPA: High School - 3rd Qt	0.071			0.072		
	(0.050)			(0.055)		
GPA: High School - 4th Qt	0.121**			-0.013		
	(0.048)			(0.056)		
1st StE - 2nd Qt		0.023			-0.001	
		(0.043)			(0.052)	
1st StE - 3rd Qt		0.010			-0.089	
		(0.054)			(0.060)	
1st StE - 4th Qt		0.077*			-0.015	
		(0.041)			(0.057)	
2nd StE - 2nd Qt			0.002			0.039
			(0.046)			(0.069)
2nd StE - 3rd Qt			0.007			-0.063
			(0.062)			(0.078)
2nd StE - 4th Qt			0.073*			-0.085
			(0.044)			(0.062)
Student TS	0.425***	0.414***	0.410***	0.323***	0.316***	0.316***
	(0.024)	(0.024)	(0.025)	(0.024)	(0.024)	(0.023)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Other tests	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2329	2383	2128	1939	2083	1701
R^2	0.618	0.622	0.607	0.500	0.511	0.505

Dependent variable: Math test score in grade 7. Standard errors in parentheses. Controls include a student's gender, year, month and place of birth, parents' place of birth, PC availability, grades in math and language, peers' attitude toward education, school and classroom share of migrants, household size, grade retention status, number of books at home, religious denomination and SDQ scores. Other tests include grade 5 tests in perception speed, logic, ICT and math or German test scores respectively. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix

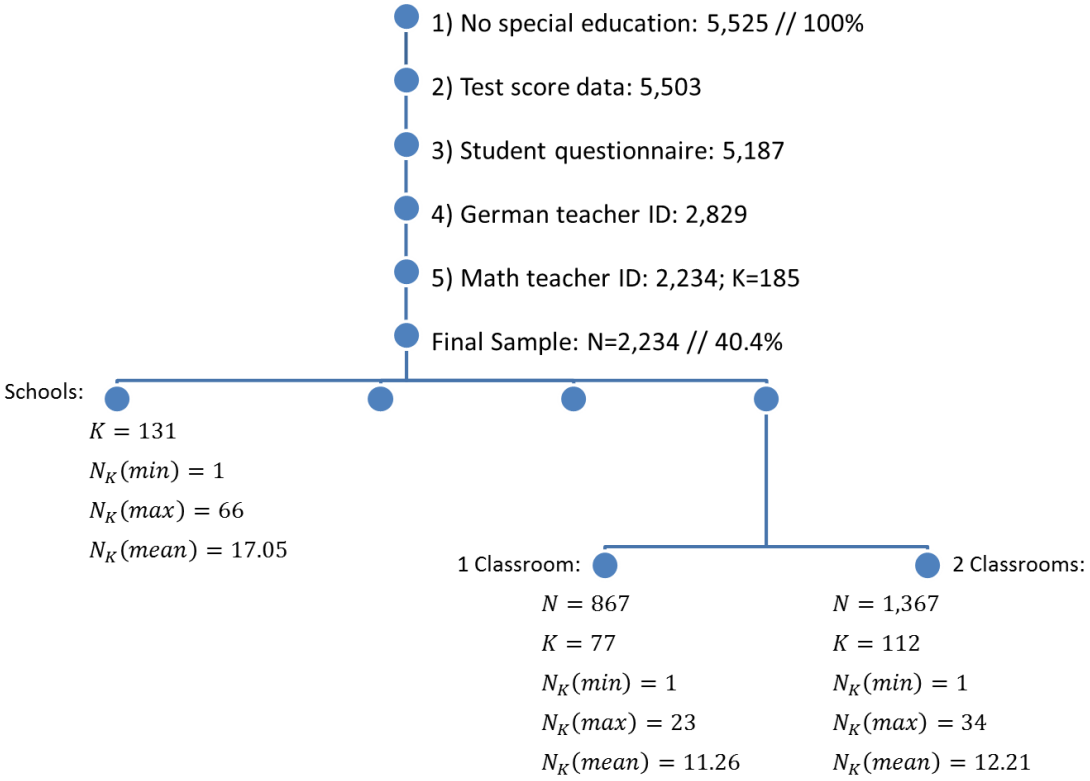

Figure A1: Matching of students' data with respective information

Table A1: Main regression results of language teachers' pre-service cognitive and pedagogical skills in 9th grade cohort of NEPS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
GPA: High School	0.098***	0.052*	0.053*							0.005
	(0.035)	(0.031)	(0.031)							(0.031)
Grade: 1st State Exam				0.045	0.009	-0.008				0.002
				(0.029)	(0.026)	(0.021)				(0.030)
Grade: 2nd State Exam							0.086***	0.021	0.021	-0.018
							(0.029)	(0.025)	(0.019)	(0.025)
L. Student TS R.Speed	0.198***	0.130***	0.068***	0.195***	0.124***	0.074***	0.203***	0.135***	0.081***	0.068***
	(0.015)	(0.013)	(0.011)	(0.014)	(0.012)	(0.011)	(0.016)	(0.013)	(0.011)	(0.012)
L. Student TS Vocabulary	0.482***	0.416***	0.147***	0.501***	0.429***	0.161***	0.491***	0.422***	0.152***	0.146***
	(0.016)	(0.018)	(0.017)	(0.014)	(0.016)	(0.016)	(0.016)	(0.017)	(0.017)	(0.020)
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes
Other tests	No	No	Yes	No	No	Yes	No	No	Yes	Yes
Observations	6734	5298	5264	7400	5796	5759	6483	5102	5073	4051
R^2	0.355	0.419	0.562	0.361	0.424	0.562	0.361	0.424	0.567	0.561

Dependent variable: German test score in the end of grade 9. Standard errors in parentheses. Controls include a student's gender, year, month and place of birth, parents' place of birth, PC availability, grades in math and language, peers' attitude toward education, school and classroom share of migrants, household size, grade retention status, number of books at home, religious denomination and SDQ scores. Other tests include grade 5 tests in perception speed, logic, ICT and math or German test scores respectively. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

ifo Working Papers

No. 242 Doerrenberg, P. and A. Peichl, Tax morale and the role of social norms and reciprocity. Evidence from a randomized survey experiment, November 2017.

No. 241 Fuest, C., A. Peichl and S. Siegloch, Do Higher Corporate Taxes Reduce Wages? Micro Evidence from Germany, September 2017.

No. 240 Ochsner, C., Dismantled once, diverged forever? A quasi-natural experiment of Red Army misdeeds in post-WWII Europe, August 2017.

No. 239 Drometer, M. and R. Méango, Electoral Cycles, Effects and U.S. Naturalization Policies, August 2017.

No. 238 Sen, S. and M.-T. von Schickfus, Will Assets be Stranded or Bailed Out? Expectations of Investors in the Face of Climate Policy, August 2017.

No. 237 Giesing, Y. and A. Music, Household behaviour in times of political change: Evidence from Egypt, July 2017.

No. 236 Hayo, B. and F. Neumeier, Explaining Central Bank Trust in an Inflation Targeting Country: The Case of the Reserve Bank of New Zealand, June 2017.

No. 235 Buettner, T. und M. Krause, Föderalismus im Wunderland: Zur Steuerautonomie bei der Grunderwerbsteuer, März 2017.

No. 234 Blesse, S. und F. Rösel, Gebietsreformen: Hoffnungen, Risiken und Alternativen, Januar 2017.

No. 233 Hayo, B. and F. Neumeier, The (In)Validity of the Ricardian Equivalence Theorem – Findings from a Representative German Population Survey, December 2016.

No. 232 Fritzsche, C. and L. Vandrei, The German Real Estate Transfer Tax: Evidence for Single-Family Home Transactions, November 2016.

No. 231 Nagl, W. and M. Weber, Stuck in a trap? Long-term unemployment under two-tier unemployment compensation schemes, November 2016.

No. 230 Neumeier, F., Do Businessmen Make Good Governors?, November 2016.

No. 229 Gutmann, J., M. Neuenkirch and F. Neumeier, Precision-Guided or Blunt? The Effects of US Economic Sanctions on Human Rights, November 2016.

No. 228 Felbermayr, G., R. Aichele, I. Heiland, A. Melchior and M. Steininger, TTIP – Potential Effects on Norway, (November 2016) revised version December 2016.

No. 227 Schueler, R., Educational inputs and economic development in end-of-nineteenth-century Prussia, October 2016.

No. 226 Riem, M., Does political uncertainty influence firm owners‘ business perceptions?, October 2016.

No. 225 Enzi, B. and B. Siegler, The Impact of the Bologna Reform on Student Outcomes – Evidence from Exogenous Variation in Regional Supply of Bachelor Programs in Germany, October 2016.

No. 224 Roesel, F., Do mergers of large local governments reduce expenditures? – Evidence from Germany using the synthetic control method, October 2016.

No. 223 Schueler, R., Centralized Monitoring, Resistance, and Reform Outcomes: Evidence from School Inspections in Prussia, October 2016.

No. 222 Battisti, M., R. Michaels and C. Park, Labor supply within the firm, October 2016.

No. 221 Riem, M., Corporate investment decisions under political uncertainty, October 2016.

No. 220 Aichele, R., I. Heiland and G. Felbermayr, TTIP and intra-European trade: boon or bane?, September 2016.

No. 219 Aichele, R., G. Felbermayr and I. Heiland, Going Deep: The Trade and Welfare Effects of TTIP Revised, July 2016.