I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 18279

The Importance of Teachers and Socioeconomic Background for Students at Risk of Dyslexia

Nicolai Kristensen

NOVEMBER 2025

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 18279

The Importance of Teachers and Socioeconomic Background for Students at Risk of Dyslexia

Nicolai Kristensen
VIVE and IZA

NOVEMBER 2025

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Importance of Teachers and Socioeconomic Background for Students at Risk of Dyslexia*

Approximately 7-10% of the population have some degree of dyslexia, and students with this disability are likely to be more dependent on qualified teaching. I analyze this tenet using population-wide Danish administrative records of public schools, where subject teachers are linked over time to classrooms and students, and identification is achieved using a within-school between-class model. I find that qualified teachers improve student outcomes mostly at the bottom of the student skill distribution and that students tested as at-risk of dyslexia in 4th grade make smaller progress in 6th grade vis-à-vis their peers, except for at-risk students from high SES schools who catch up with their peers. Unqualified teaching in combination with low SES schools impedes the ability of dyslexic students to reach their potential, and upholds the inter-generational correlation in education.

JEL Classification: I21, I24

Keywords: reading ability, teacher quality, socioeconomic status, dyslexia

Corresponding author:
Nicolai Kristensen
VIVE - The Danish Centre of Applied Social Science
Herluf Trolles Gade 11
1052 Copenhagen K
Denmark
E-mail: nikr@vive.dk

* Financial support from the Danish National Research Council (grant number 8091-00042B) is gratefully acknowledged. I thank Paul Bingley, Jens Dietrichson, Rasmus Landersø and Hans Henrik Sievertsen for useful comments, Nils Djervad for useful discussions, and Johannes Rosendal for excellent research assistance.

1 Introduction

Many countries lack qualified public school teachers and the shortage has been increasing over many years. OECD (2020) reports that in 42 countries, students attending socio-economically disadvantaged schools were exposed to more shortages of educated staff than their peers in advantaged schools.[1]

In addition, teacher sorting, whereby high-achieving students are taught by the best-qualified teachers, is pervasive in many countries. Teacher sorting has been studied intensely in previous research (*i.a.*, Rivkin et al. (2005); Boyd et al. (2005); Aaronson et al. (2007) Jackson (2009); Clotfelter et al. (2010); Hanushek and Rivkin (2010); Bonhomme et al. (2016); and the systematic review of 120 studies by Nguyen et al. (2020)), and these many well-crafted studies jointly demonstrate how the best teachers sort into schools populated by pupils from affluent homes of high socioeconomic status (SES).[2]

The problem of teacher sorting and shortage of qualified public school teachers is further amplified if the consequences in terms of teaching and student learning outcomes primarily are borne by vulnerable students such as students at risk of being dyslexic and students from low SES areas.

There are at least two reasons why students with low reading ability (including students at risk of being dyslexic) deserve special attention in studies of teacher qualifications. Firstly, about 7-10 percent of a population are dyslexic and poor readers

[1] In the United States there are reportedly at least 55,000 vacant teacher positions and 270,000 positions filled by under-qualified staff. (See https://www.teachershortages.com/.) For the United Kingdom Sibieta (2020) found that about 20 percent of schools in the least disadvantaged areas had teacher vacancies while the corresponding number was about 30 percent of schools in the most disadvantaged areas outside of London, and about 46 percent of schools in disadvantaged areas in London. For Denmark, about 15-20 percent of all public school teachers do not have a degree from teacher college, c.f. AE Rådet (2023).

[2] Generally, the studies have found statistically significant positive effects of teacher quality measured by teacher-college education or other quality indicators on student test outcomes. Another general finding in these studies is that the highest impact of teacher quality on student outcomes is found for disadvantaged groups such as students from low SES families.

therefore constitute a large and important group of students with special needs.[3] In general, poor readers are likely to benefit more than average from a teacher who understands and caters for each individual students reading disability, see Taylor et al. (2010), and a large proportion of these students are likely dyslexic. Like other learning difficulties, dyslexia demands specialized teaching and high didactic skills, see Solheim et al. (2024).

Secondly, while parents' involvement in their children's education is important in general, it plays a crucial role for students with special needs, such as students with reading difficulties. Parents in high SES families are generally found to assist their children more with their homework (Guryan et al. (2008), Kalil et al. (2012), Falk et al. (2021)), while children from low SES families become more dependent on teacher qualifications. Problems with poor reading ability are, therefore, likely to manifest themselves more strongly in combination with low SES family background. As dyslexia oftentimes is inherited – and one or both parents therefore potentially challenged in reading themselves – the importance of qualified teachers is amplified in the case of dyslexia.[4]

This paper is the first to estimate causal effects of teacher quality on reading ability across the student skill distribution. The paper is based on data of exceptional richness. Administrative individual level panel data makes it possible to link population-wide public school teachers to students at the school, year, grade, subject and class level. Concerning teachers, the data include detailed information about each public school teacher, such as formal teacher education (teacher college), the teachers' grade point average (GPA) from teacher college, and GPA from high school of those with no teacher-

[3] *Sources*: The Danish Library and Expertise Center for people with print disabilities, https://nota.dk/om-nota/english and The British Dyslexia Association, https://www.bdadyslexia.org.uk/dyslexia.

[4] Snowling and Melby-Lervåg (2016) estimate that if a child has a parent with dyslexia, the probability of the child having dyslexia is on average 45%.

college education. Information about students includes, *i.a.*, diagnoses of students with attention deficit disorder (ADD) both for the individual student and computed also for classroom peers, ethnic background, parental education levels, and family income. Reading ability is measured in National Tests in Reading. These tests compose of scores in each of three domains: language comprehension, text comprehension and decoding. The score on decoding is a very good proxy for a test of dyslexia, where a low decoding score strongly signals that the student might be dyslexic. An at-risk of dyslexia indicator is coded as one if the decoding score is in the lowest 10 percent, and zero otherwise.

The paper has two key findings. Firstly, qualified teachers are found to improve outcomes mostly for students at the bottom of the student skill distribution. One interpretation of this finding (outlined in more detail below) is that college educated teachers appear better at individualized teaching differentiation, whereby the teacher adapts method, pace and evaluation of the teaching to the individual student's needs. Secondly, being at-risk of dyslexia in grade 4 is generally a sign of lower progression in grade 6, but at-risk students attending high SES schools progress more than their student peers at the same school. This difference in reading progress of at-risk dyslexic students along the school SES gradient is important as it indicates that there is an unsolved potential for dyslexic students from the lower parts of the SES distribution.

From a modeling perspective, this paper also makes a methodological contribution. A common identification strategy for teacher effects on students test outcomes is to estimate effects by comparing outcomes between classes at the same school, year, grade, and subject. If class A, say, has a college educated teacher and class B does not, this difference in teacher qualifications drives the parameter estimate of having a qualified teacher. In the methodology-section it is outlined how this identification strategy implicitly rests on a linearity assumption, which may fail due to teacher sorting both among college educated teachers *and* sorting among teachers with no formal teacher credentials. Schools in high SES areas are able to attract better non-qualified teachers than schools in low SES areas.

This study has ties to two major strands of literature. First, it follows in the wake of the many studies on the importance of teacher quality and teacher education on student outcomes and the large literature documenting how teacher sorting results in low-achieving students often being taught by the least-qualified teachers.[5] Teacher sorting underscores the importance of a strong identification strategy in order to claim causality. Secondly, part of the motivation behind this study are the underlying ties to inequality in education, and inter-generational inheritance in educational outcomes. Previous studies in this strand of literature include, *i.a.*, Black et al. (2005), Black and Devereux (2011), and Björklund and Salvanes (2011). The findings in this paper provides an additional path for explaining the high inter-generational correlation in education.

The paper proceeds as follows. In the next sections, I describe data and methodology, including a description of the identification strategy. This is followed by a presentation of results, discussion, and concluding remarks. In the appendix, I describe an alternative identification strategy that builds on within class variation, and shows results based on this approach.

2 Data

The paper is based on administrative registry data from Statistics Denmark. Each individual has his or her own central personal register (CPR) number, which makes it possible to link information across many registries and over time. The registries cover

[5]Including the many citations provided above. Based on Danish data, Gensowski et al. (2024) also find evidence of teacher sorting in a Danish context.

the entire Danish population and include very detailed information on demographics, education, and labor market relations.

Especially relevant for this study, all public school teachers can be linked over time to schools, grades, classes, and subjects they teach for core subjects under the so-called National Tests where students are tested in Reading (Danish) in grades 2, 4, 6, and 8.[6]

Furthermore, the educational background of all public school teachers is observed, including whether they have a degree from a teacher college and their grade point average from teacher college.[7] For public school teachers with no formal degree from teacher college the high school GPA is observed. Lastly, the work history is known in detail for all teachers irrespective of educational background.

2.1 The National Test Data

Reading scores come from National Tests in Reading. The tests composes of a score in each of the following three domains: language comprehension, text comprehension and decoding. Jointly, these three scores measure the students ability in Danish with an emphasis on reading (as opposed to writing). The score on decoding is a measure that serves as a good proxy for a test of dyslexia (Elbro and Kristensen (2025)), and the decoding scores in grades 2 and 4 can therefore serve as early indicators of which students are at-risk of being dyslexic. A low decoding score strongly signals that the student might be dyslexic. An at-risk indicator is coded as one if the decoding score is in the lowest 10 percent, and zero otherwise.[8]

The National test outcomes originates from a Rasch-calibrated logit scale for each

[6]Tests in mathematics occur in grades 3 and 6. Test results also exist for English (Science) in grade 7 (8). In this study, only the Reading scores are used.

[7]Grade point average from teacher college is available for teachers completing college in 2004 or later.

[8]The threshold for being "at risk of dyslexia" is set at 10% (nationwide) as this aligns with the approximately 7-10% of a population being dyslexic.

of the three domains. For each domain, the test score is standardized to mean zero and standard deviation one within each year. Subsequently, for each student and domain, the average test score is computed across the three domains and this average is once again standardized to mean zero and standard deviation one within each year.[9] I use this combined standardized test score as the student outcome measure in 6^{th} grade, and also include the lagged outcomes from 2^{nd} and 4^{th} grade as control variables. The measure for being "At risk" of dyslexia is based only on the decoding score in 2^{nd} and 4^{th} grade. The choice of analyzing the teacher impact on students in 6^{th} grade is made since it is important to control for lagged student test scores, and test scores of 2^{nd} grade students are much less stable and therefore provides a more noisy measure of student ability than test scores in 4^{th} grade.

There are missing observations in these scores when a student for one reason or another did not take the test. Some of the non-tested students are exempted from the national tests because they face severe difficulties and the test therefore becomes meaningless and unnecessary.[10] Students who did not participate in the test in, say, 2^{nd} grade Danish, but who took the test in 4^{th} grade are included in the analysis. Since it is important for parameter identification that we can condition on previous individual outcomes, students who do not take the test in Danish in 2^{nd} grade nor 4^{th} grade – in part likely to be the lowest-qualified – are excluded from the analysis.

2.2 Measurement of Teacher Quality

I make use of two measures of teacher quality. The first measure is a binary indicator of whether the teacher has completed teacher college or not. It is a key indicator with

[9]This follows standard procedure and has previously been adapted, e.g., by Beuchert and Nandrup (2017).

[10]The purpose of the test is for the teacher to obtain knowledge about the status of the student. For the very most poor readers the teacher (and the parents) will have observed the difficulties.

direct link to policy and the above-mentioned increasing lack of certified teachers in many countries. In Denmark, people with no formal teaching background can enter the teacher profession and, next to teaching, take courses to achieve a certificate as a "merit teacher". We therefore also include an indicator for this type of teacher qualification.

A second measure of teacher quality is the GPA from teacher college. Arguably, this is a very good measure of the mix of skills a good teacher should possess, including scores on classroom leadership, building teacher-student relations, and didactic skills. GPA from teacher college exists for all teachers who completed teacher college in 2004 or later.[11]

An alternative to these two simple measures of teacher quality could be to adapt a teacher value added (VA) approach. However, the student-teacher link can only be made over few years and a VA estimate will therefore likely be highly volatile.[12] Furthermore, Guarino et al. (2015) simulate data under a variety of scenarios to study the VA approach' ability to produce accurate measures of teacher effects. They conclude that the risk of misclassifying teachers as high- or low-achieving can be substantial. With this caveat, and given our relatively short panel-length, I choose to focus on whether teachers completed teacher college or not and their GPA from college as measures of teacher quality. Policy-wise, it may also be easier to translate results based on completed teacher college than results based on a more complex measure such as Value Added. This consideration becomes especially relevant given the high and increasing share of public school teachers with no teacher college certificate.

[11]For the "merit teachers" and the teachers with no certificate from teacher college we do not observe a GPA from teacher college. For this group, we use their high school GPA. Admittedly, this measure is not likely to measure teacher qualifications nearly as well as the teacher college GPA, but it may nevertheless proxy for the teachers cognitive skills and zeal for work. These continuous GPA measures of teacher quality make it possible to adhere to within subject and class identification as further explained in the method section.

[12]Early Value Added (VA) studies include Hanushek (1971). Later studies include Rivkin et al. (2005) and Aaronson et al. (2007), and the influential method refinements made by Chetty et al. (2014).

3 Descriptive Facts

The school average parental income is computed for all parents to students in 6^{th}-grade over the school years 2014/2015 up to 2018/2019.[13] Based on the school average parental income measure I compute both school SES deciles and quartiles and use these in the analysis.

The school SES quartiles are applied to split the sample in three (Q1=bottom quartile, Q2-Q3, and Q4=top quartile) for which student, family, and class characteristics are shown in Table 1. The average annual household income is just below 80,000 euros for schools in the bottom SES quartile, rising to almost 94,000 euros in the middle half, and more than 128,000 euros in the top SES quartile.

Note from Table 1 how being diagnosed with Attention Deficit Disorder (ADD) in 6^{th} grade occurs to 1.4% - 1.8% of students with only a modest decrease as one moves up the school SES quartiles.[14] The bottom panel of Table 1 shows shares of school classes with zero, one, or more than one ADD-student in the class and this is also shown to be relatively constant across school SES quartiles.

[13] In real 2023 euros. For each student the sum of parental income enters the computation of the school average parental income, which is our measure of school SES. No distinction is made between parents living together or apart. Income of any new partner is not included.

[14] The individual diagnosis of hyperkinetic disorders in Danish children follows the WHO International Classification of Diseases, ICD-10. It is used as a standard by which individual students with ADD (including ADHD) are registered under 'F90 Hyperkinetic disorders'. This group comprises F90.0 Disturbance of activity and attention, F90.1 Hyperkinetic conduct disorder, F90.8 Other hyperkinetic disorders, and F90.9 Hyperkinetic disorder, unspecified. The measure is a simple presence or absence of such a diagnosis.

Table 1: Student, Family, and Class characteristics, by schools SES decile

	SES Quartile		
	Q1	Q2-Q3	Q4
Family			
HH income (euros)	79,324	93,744	128,336
Mother Non-West Immi	21.1%	10.7%	8.8%
Father Non-West Immi	20.2%	9.7%	7.9%
Mother years education	13.6	14.6	15.5
Father years education	13.5	14.3	15.4
Student			
ADD	1.8%	1.8%	1.4%
At-Risk 2nd grade	10.9%	7.2%	4.6%
At-Risk 4th grade	12.0%	8.0%	4.9%
Test score in Danish 2nd grade	-0.20	0.04	0.24
Test score in Danish 4th grade	-0.20	0.04	0.27
Dyslex tested	5.4%	5.1%	4.1%
Dyslexic	66.6%	62.4%	48.8%
Borderline dyslexic	25.9%	28.0%	34.8%
Not dyslexic	7.6%	9.6%	16.4%
Class			
0 ADD	70.7%	69.5%	72.2%
1 ADD	23.4%	24.2%	23.3%
2+ ADD	6.0%	6.2%	4.5%

Note: Students are observed in 6^{th} grade. Total number of students included here is 231,160 over the school years 2014/2015 to 2018/2019 and across 1,179 public schools. School SES quartiles are based on average parental income at the school. Quartiles 2 and 3 are combined.

Table 1 also shows how the share of parents with non-western immigrant background is above 20% in school SES bottom quartile while about 9-10% in the other SES quartiles. The high share with non-western family background in the Q1-column of Table 1 may be one explanation why students from the bottom school SES quartile have much lower average test scores in Danish in second and fourth grade than the students from SES quartiles 2-3 and quartile 4. In part, the low Danish scores can be a manifestation of foreign maternal (paternal) language among some of these students possibly in combination with dyslexia.[15]

While the national test is taken by students in general, only selected students undertake a special dyslexia test, Elbro and Kristensen (2025). The share of students actually tested for dyslexia is almost constant across school SES deciles even though the share at-risk of dyslexia decreases linearly across deciles, see Figure 1. These striking differences between share at-risk and share of dyslexia tested across the school SES gradient arise in part because parents have a legal right to claim such a test once during their child's compulsory school life, and high SES parents are likely much more aware of this right. Also, teachers at high SES schools may be more attentive to the relevance of a dyslexia test.

Teacher qualifications differ across the school SES deciles. The best-qualified teachers, as measured by GPA, appear to sort (or be sorted) towards schools where the students on average come from affluent homes. In Figure 2, the grade point average (GPA) from teacher-college and the GPA from high school for public school teachers without a degree from teacher-college, are compared for school deciles 1 and 10.[16]

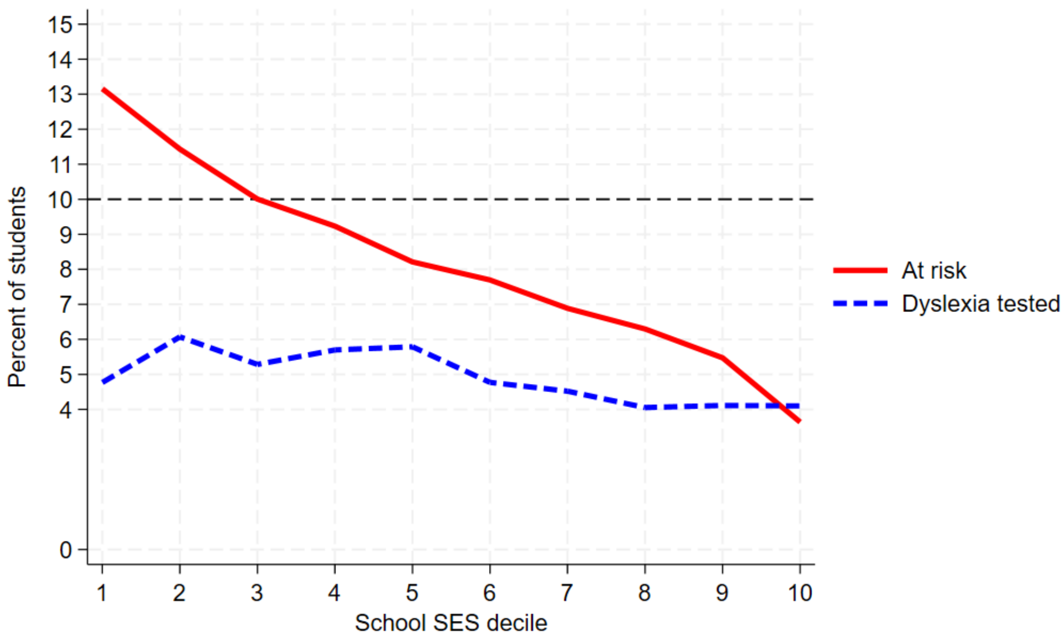
Panel (c) of Figure 2 shows how schools in the top decile of parental income are able to attract teachers with relatively high GPAs from high school while the distribution of high school GPA of teachers at schools in the bottom decile are much lower and the spike around the modus is even below the minimum GPA found in schools in the top decile (decile 10).

Differences in the GPA distribution between school decile 1 and 10 are also clear

[15] While reading difficulties due to decoding is a sign of dyslexia other impediments to learn reading can be understanding the text. This is a classic problem for people with limited vocabulary and low understanding of the language, which often will be the case for people reading in a language other than their maternal language such as immigrants. Gellert (2009) analyze the group challenged in both dimensions. Elbro and Kristensen (2025) find that students with immigrant family background are tested significantly less often for dyslexia given their decoding score in the national test.

[16] Since Danish (i.e. reading, writing and spelling) is a core subject we expect the teachers with no teacher-college qualifications to be positively selected within each school, and hence that they likely perform better than would other teachers with no formal degree.

Figure 1: Share at-risk of being Dyslexic and Share Dyslexia Tested, across school SES deciles



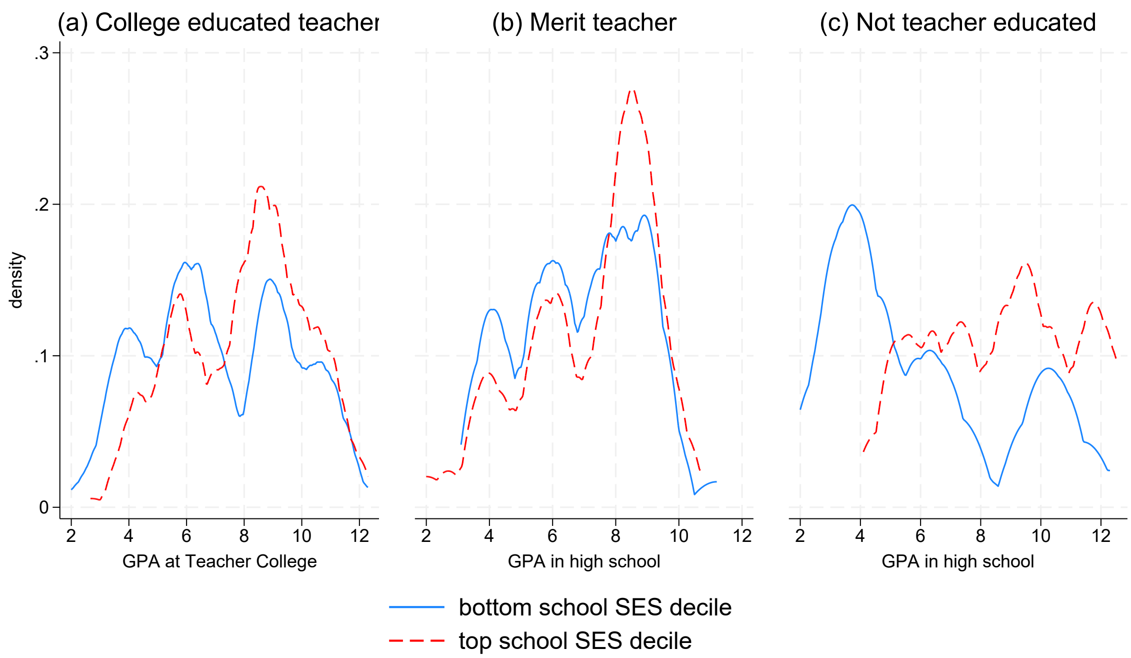
Note: Students observed in SES deciles on the horizontal axis. Being at-risk is defined as the group of students with a decoding score in the lowest 10 % nationwide.

from panel (a) and panel (b) of Figure 2. However, while clear when we compare the top and bottom ten percent the differences are much less pronounced and not necessarily linear across all 10 deciles, see Table 6 in appendix B. A basic yet key point to note is that teachers with no credentials also are very heterogeneous and also differ across the school SES distribution. In other words, teacher sorting is also present among teachers with no formal teacher qualifications.

4 Methodology

Studies of the impact of teacher quality on student outcomes generally need to address the challenge to causal analysis that arises from teacher sorting. Identification of causal effects of teacher quality on students' test outcomes becomes difficult because teachers

Figure 2: Teacher-college GPA or high school GPA, by type of teacher background and by deciles of the school-level socioeconomic background



Note: The figure is based on GPAs of teachers in Danish in grade 6 over the school years 2014/2015 to 2018/2019.

and students are not randomly matched, and the choice of method needs to be able to deal with such systematic sorting. This problem is well-known in the literature and several econometric strategies have been applied to deal with teacher-pupil sorting, see e.g., Barrow & Rouse (2007).

4.1 Within school, grade, and subject and between classes identification[17]

First, we use the variation in teacher quality measures for teachers at the same school, year, grade, and subject but in different classes at the school. With school and year fixed effects any causal impact of observable teacher characteristics is identified if students

[17]See Andersen et al. (2024) for a recent study that uses a similar identification strategy.

are randomly allocated to the classes, conditional on school and year.

By comparing teachers employed at the same school we (seemingly) purge the problem of sorting. However, this is the case only under an implicit underlying assumption of linearity in the impact of teacher quality over the distribution of schools. To be clear, let us compare the within school variation in teacher quality in a low SES school with a high SES school.

In *low SES schools*, on average, we compare student test scores in a class with a college educated teacher with relatively low college GPA to student test scores in another class at the same school, subject, grade, and year but where the teacher is not college educated. This non-educated teacher will, on average, have relatively low GPA from high school.

The opposite holds true for *high SES schools* where, on average, we compare teacher college educated teachers with high GPA with non-educated teachers with high high-school GPA. Unless the parameter estimate for a college educated teacher vis-à-vis a non-educated teacher is the same across GPA levels, we will get biased estimates if only one single parameter is allowed for.

This underlying linearity assumption can be tested simply by including interaction terms between school SES deciles and teacher quality and such a test is carried out in the empirical part of the paper, where the average parental income is used at schools in quartiles 2 and 3 (25-75 percentiles) as reference group and allow for different parameters for the bottom (0-25) and top quartile (75-100). The basic but important point to note (once again) is that teacher sorting also arises among non-college educated teachers.

The simplest version of the econometric model specification for this type of model can be written as

$$TestScore_{itcs} = \beta_0 + \beta_1 TC_c + \gamma_s + \theta_t + \varepsilon_{itcs}, \tag{1}$$

where $itcs$ are indices for individual i, time t, class c, and school s. The γ_s and θ_t signify school and year fixed effects, respectively, and ε_{itcs} is a random idiosyncratic error term. The parameter of interest is β_1, which gives the causal impact of teacher college education (TC_c) (with no teacher credentials as the reference group) on student test score ($TestScore_{itcs}$).

A more elaborate version of model (1) includes controls for student, parents, class and teacher characteristics. In order to test for the implicit linearity assumption mentioned above, we interact indicators for school SES deciles with teacher quality captured by the β_6^q and β_7^q - parameters in equation (2) below, i.e., with quartiles (2 and 3) as a joint reference group.

$$\begin{aligned} TestScore_{icts} = \beta_0 + \beta_1 TC_c + \beta_2 AtRisk4_i + \\ \beta_3 \left(AtRisk4_i \cdot TC_c\right) + X_i\beta_4 + \left(AtRisk4_i \cdot X_i\right)\beta_5 + \\ \left(School - SESquartile_{q1} \cdot TC_c\right)\beta_6 + \\ \left(School - SESquartile_{q4} \cdot TC_c\right)\beta_7 + \\ \gamma_s + \theta_t + \varepsilon_{itcs} \end{aligned} \tag{2}$$

Equation (2) includes the possibility of an interaction term between teacher college education (TC_c) and an indicator for the student being in the group who scored among the lowest 10 percent in decoding in 4^{th} grade Danish test (and hence the indicator "At-Risk4" equals 1).

The parameters of main interest are β_1, β_3, β_6 and β_7. The β_1-parameter gives the effect measured in percent of a one SD of having a teacher with a college education (with no education as reference) in Danish in 6^{th} grade, and $\beta_1 + \beta_3$ can be interpreted as the effect of having a teacher with teacher college education in 6^{th} grade for a student at-risk of dyslexia measured by his or her 4^{th} grade decoding score. β_6 and β_7 allow for non-linearity across school SES levels. The base are the two middle school SES quartiles (quartiles 2 and 3), while β_6 yields an estimate for the lowest achool SES quartile and β_7 yields an estimate for the highest school SES quartile. A test of non-linearity simply is whether one or both of β_6 and β_7 are statistically significantly different from zero.

Using model specification (2), identification is challenged if students can sort into classrooms within the school. In a Danish context, this would be a rare event.[18] In particular, students and their parents would not be likely to induce any such sorting, but the school leader, possibly in collaboration with teachers in a given school, year, and grade, could possibly allocate a student to a selected teacher. Even if this would occur, we do condition on lagged outcomes in terms of standardized test score results in Danish grade 2 and 4, which further strengthen identification. As mentioned by Aaronson et al. (2007), lagged test scores also account for the cumulative inputs of prior years while allowing for a flexible autoregressive relationship in test scores. Conditioning on past test scores means that we implicitly control for time-constant family background. It is reasonable to conjecture that time-varying student characteristics, such as parental divorce, may influence student test scores. But given our statistical model, bias is only introduced to the teacher college-education estimate if students are assigned to teachers *based* on these unobservable changes, which appears unlikely. In addition, we include a number of important observable student, family, and class characteristics because they may be correlated with behavioral changes that influence achievement and may account for group differences in gain trajectories.

The second identification approach uses within class variation. This approach is used here for sensitivity / robustness analysis, and both the method and results are

[18]Kalogrides and Loeb (2013) review the research on within school sorting and mention Morgan and McPartland (1981), Clotfelter et al. (2002), and Conger (2005) as three studies examining within-school sorting by race using data from a large number of schools and students. The three studies reach the same conclusion: classroom segregation is higher in high schools and middle schools than in elementary schools.

described in the appendix material, see Section A.

4.2 Unconditional Quantile Regression

Model (2) and (3) are estimated using fixed effects ordinary least squares (OLS), which yields estimates of the conditional mean.

Given our special interest in the lower end of the distribution of reading skills, the unconditional quantile regression method developed by Firpo et al. (2009) includes fixed effects. One advantage of the unconditional quantile regression model is that the quantiles are defined pre-regression, and, therefore, the model is not influenced by any right-hand-side variables (Killewald and Bearak (2014)).[19] In unconditional quantile regression, including fixed effects to adjust for selection bias is done without redefining the quantiles. Focus is on whether or not there is a 'multiplier effect' whereby early reading problems, found in second or fourth grade national tests, are aggravated when combined with relatively low teacher quality and/or parents with low SES background, in which case additional support from the teacher may be required.

These unconditional quantile estimates also rely on the within school between class identification strategy.

[19] The right-hand-side variables help control for relevant background factors, but their inclusion has no effect on which observations are defined to be at a given quantile of the student test score distribution.

5 Results

5.1 Estimates of Average Teacher College Effects on Students Reading scores

Results from estimating model (2) with increasingly many controls, including lagged outcomes, are shown in Table 2.

Table 2: Within school between class model

	(1)	(2)	(3)	(4)	(5)
Teacher coll. educ.	0.00766***	0.00489***	0.00376***	0.00400***	0.00388***
	(0.000499)	(0.000500)	(0.000651)	(0.000428)	(0.000429)
AtRisk4		-0.759***	-0.770***	-0.0186***	-0.0441***
		(0.000956)	(0.00253)	(0.00342)	(0.00399)
T.coll.educ. x At-Risk4			0.0143***	-0.000361	-0.000431
			(0.00317)	(0.00351)	(0.00367)
Test score grade 4				0.421***	0.422***
				(0.000818)	(0.000816)
Test score grade 2				0.188***	0.188***
				(0.000315)	(0.000308)
School fixed effects	X	X	X	X	X
Time fixed effects	X	X	X	X	X
Parental controls	X	X	X	X	X
Class controls	X	X	X	X	X
Student controls	X	X	X	X	X
Interaction terms (β_5)					X
Observations	229,652	229,652	229,652	229,652	229,652

Note: Dependent variable is combined standardized test score in Danish in grade 6. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Control variables further described in Table B in the supplementary material.

The top row shows the parameter estimate for having a teacher-college educated teacher in Reading (Danish) in 6^{th} grade. The results change some as we add more controls although they remain relatively robust. Not surprisingly, the parameter estimates for the lagged test outcomes ("Test score grade 4" and "Test score grade 2") are high and very significant. This underscores the need to control for these lagged outcomes for the identification strategy to be credible even when one compares students in the same school, year, grade, and subject. The standardized test scores increase by 0.4 percent of an SD if the teacher is college-educated (top row). While highly significant it remains a small effect. The total effect of teacher-college for students at-risk of dyslexia (row 1 + row 3) is not statistically significantly different from the average effect of a teacher-college educated teacher and the point estimate in row 3 (teacher college educated teacher interacted with students at risk of dyslexia) is very close to zero. On average, students at risk of dyslexia score lower than their peers even when we control for past test scores in both second and fourth grades together with many other controls. This suggests that dyslexic students do not meet sufficient measures at the school to counter their disability. Given relatively low test outcomes in 4^{th} grade one could imagine the students at-risk would catch up (so that the indicator had a positive sign), but on average this is not the case.

5.1.1 The Impact of College Educated Teachers across School SES Levels

Special attention is here given to the gradient in teacher college background across schools with different SES profiles. To this end, I split schools SES into a bottom and top quarter and keep the two middle quarters joint as reference group. The estimated model is equation 2, see column (2) of Table 3.

Both the interaction parameters, β_6 and β_7 of equation (2), are estimated to be significantly positive. The highest impact of having a teacher college educated teacher is found for the low SES schools ($\beta_1 + \beta_6$). The effect is highly non-linear as β_7, the interaction parameter for top quartile, also is positive and significant. It is not clear

Table 3: Testing for non-linearity

	(1)	(2)
Teacher coll. educ. (β_1)	0.00388***	0.00244***
	(0.000429)	(0.000503)
School low-SES x Tcoll (β_6)		0.00639***
		(0.00137)
School high-SES x Tcoll (β_7)		0.00315*
		(0.00136)
School fixed effects	X	X
Time fixed effects	X	X
Parental controls	X	X
Class controls	X	X
Student controls	X	X
Interaction terms (β_5)	X	X
Observations	229,652	229,652

Note: Dependent variable is combined standardized test score in Danish in grade 6. Within school between class model using school SES interactions. To ease comparison, column (1) includes the results from Table 4, column (4). Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

whether the nonlinearity is driven by sorting among college educated teachers or sorting among the non-college educated teachers (or both).

In order to further investigate how teachers affect at-risk students test scores across different school SES levels, I split the sample by school SES quartiles (q1, q2-3 and q4) in order to obtain separate estimates for teacher impact on at-risk students across school SES quartiles, see Table 4. Columns (2)-(4) show sub-sample estimates across school SES quartiles.

Several striking findings appear. First, the key parameter estimates differ significantly across levels of school SES. For the top 25% most affluent schools (measured by average parental income), see column (4), being at-risk of dyslexia in grade 4 has a *positive* and relatively large impact on the students outcome in grade 6. For lower levels of school SES, Q1 and Q2-3 in columns (2) and (3), it is also relatively large but

Table 4: Samples split by school SES quartiles

	(1) All	(2) Q1	(3) Q2-3	(4) Q4
Teacher coll. educ.	0.00388***	0.00762***	-0.00245***	0.00978***
	(0.000429)	(0.00101)	(0.000513)	(0.000717)
Test score grade 2	0.188***	0.182***	0.186***	0.191***
	(0.000308)	(0.000500)	(0.000745)	(0.000352)
Test score grade 4	0.422***	0.403***	0.420***	0.435***
	(0.000816)	(0.000533)	(0.00101)	(0.000584)
T.coll.educ. x At-Risk4	-0.000431	-0.00915	0.0361***	-0.106***
	(0.00367)	(0.00514)	(0.00520)	(0.00430)
AtRisk4	-0.0441***	-0.0895***	-0.0581***	0.0884***
	(0.00399)	(0.00483)	(0.00502)	(0.00330)
School fixed effects	X	X	X	X
Time fixed effects	X	X	X	X
Parental controls	X	X	X	X
Class controls	X	X	X	X
Student controls	X	X	X	X
Interaction terms (β_5)	X	X	X	X
Observations	229,652	56,669	115,459	57,307

Note: Dependent variable is combined standardized test score in Danish in grade 6. Based on the within school between class model. Column (5) from Table 4 reappears as column (1). Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

with opposite (negative) sign.

This result likely links to the fact that high SES schools undertake measures to assist dyslexic students much more, including testing them for dyslexia, to subsequently offer special assistance and separate dyslexia teaching assistance. Unfortunately, information about separate dyslexia teaching assistance is not available in the data, but it is likely an important part of the underlying explanation.[20] A second likely explanation, also

[20]Elbro and Kristensen (2025) find a strong school SES gradient concerning who is tested for dyslexia, which is also illustrated here in Figure 1.

not observed in data, is that parents in these high SES schools allocate more of their own time and/or privately pay for extra dyslexia teaching. I return to this finding in the discussion, Section 6 below.

The parameter estimates for having a teacher-college educated teacher is on the other hand not important for the students from high SES schools. This may in part reflect that the teachers replacing the college educated teachers on average have a relatively strong cognitive background. It may also be that the students in high SES schools more often are taken out of the class to participate in small-group dyslexia-classes. This does take place in many schools with different intensity and without being observed in the data.

For students from low SES schools (Q1), column (2) of Table 4, being at risk of dyslexia in grade 4 is a strong predictor for lower reading scores in grade 6, even though lagged test scores are included among the control variables. Having a teacher-college educated teacher has a positive and strongly significant impact on Q1-students in general but not particularly so for those at risk of dyslexia. Lastly, for the schools in the middle of the SES-distribution, Q2-3 in column (3) of Table 4, there is a net positive impact of having a teacher college educated teacher among the students at risk of dyslexia, but not for the remaining students where the point estimate even is significantly negative.[21]

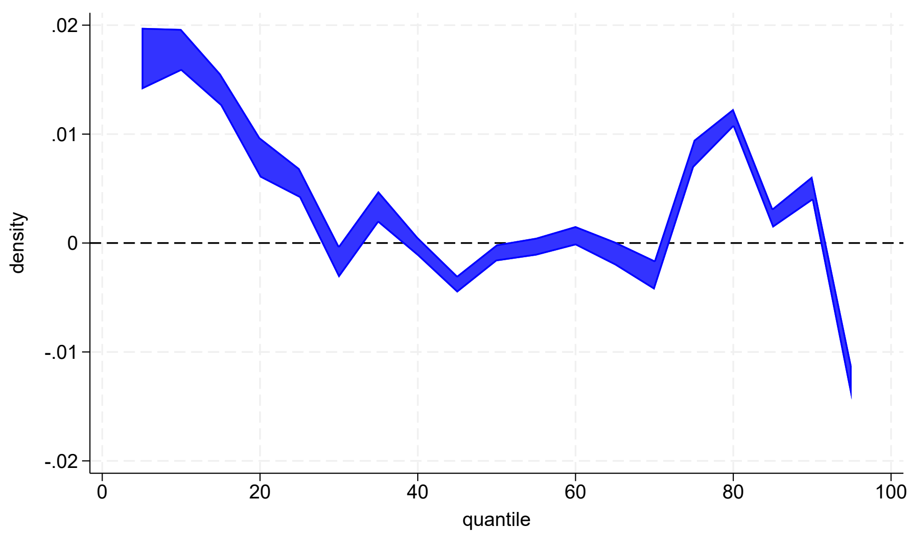
5.2 Estimates of the Distribution of Teacher Quality Effects

Average estimates may very well reflect great differences across the distribution and it is therefore relevant to estimate unconditional quantile regressions using the specification given in column (5) of Table 2.

[21] A possible, but untestable, explanation for this finding is that it may reflect a stronger focus among college educated teachers on low-achieving students at the expense of high-ability students.

Across students, teacher-college educated teachers primarily improve test scores of students in the bottom 30% of the test score distribution with point estimates of about 2 percent of an SD at the bottom of the distribution, see Figure 3. For quantiles 30-70, the Teacher-college parameter fluctuates around zero, while for quantiles 70-90 teacher college educated teachers again are found to have a positive impact (around 1 percent of a SD). This is consistent with college-educated teachers being better than teachers with no teacher-college background in implementing teaching differentiation.[22] At the very top (90^{th} percentile and above), the estimates for teacher college turn negative. Possibly, teaching differentiation sometimes happens at the expense of the very highest-achieving students).

Figure 3: Unconditional quantile regression estimates for effect of teacher college education



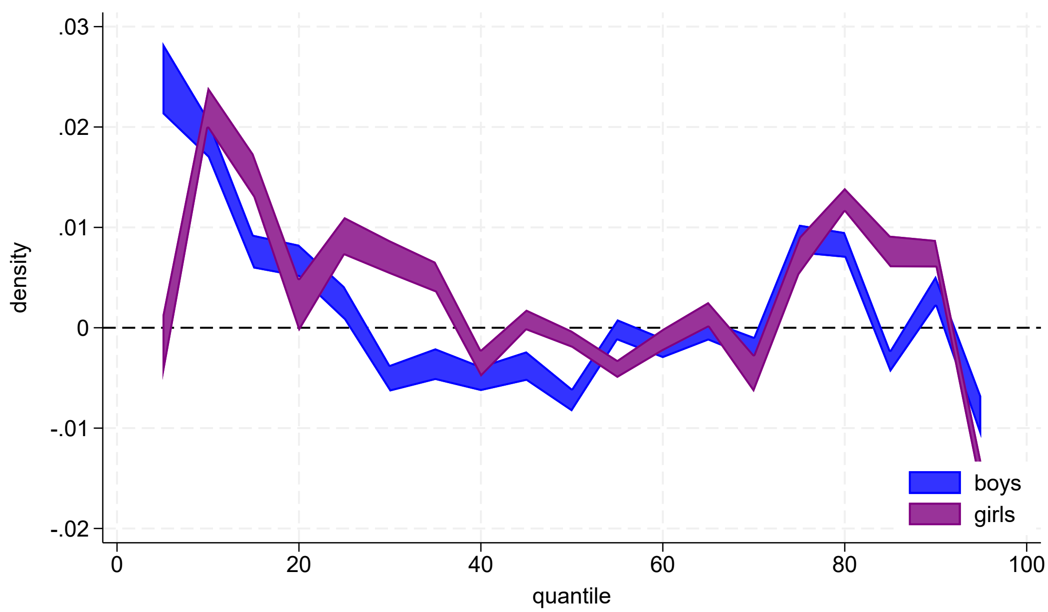
Note: Dependent variable is combined standardized test score in Danish in grade 6. 95% confidence intervals. Based on the with-in school between-class model.

[22]Teaching differentiation in primary school is a pedagogical principle that involves adapting the content, method, pace, and evaluation of the teaching to the individual student's needs and prerequisites in order to achieve common learning goals. This principle was made central in the primary school with the Primary School Act of 1993. This means that as a teacher you must know the students' different skill levels, learning styles and social prerequisites in order to create inclusive teaching that challenges and develops all students.

5.3 Heterogeneity by Gender and Class Size across the Distribution

Splitting the sample by gender, we see that boys and girls gain almost the same from having a teacher-college educated teacher although quantile estimates at the very bottom of the distribution are insignificant for girls while about 2.5% of an SD (and significant) for boys, see Figure 4. For boys, the unconditional quantile estimates yields *negative* teacher effects from 30^{th}-50^{th} percentile. This indicates that the teachers incrased attention on the poorest readers comes with a (minor) cost paid by the boys that are above the lowest echelons but below the average. The negative estimates are significantly different from zero.

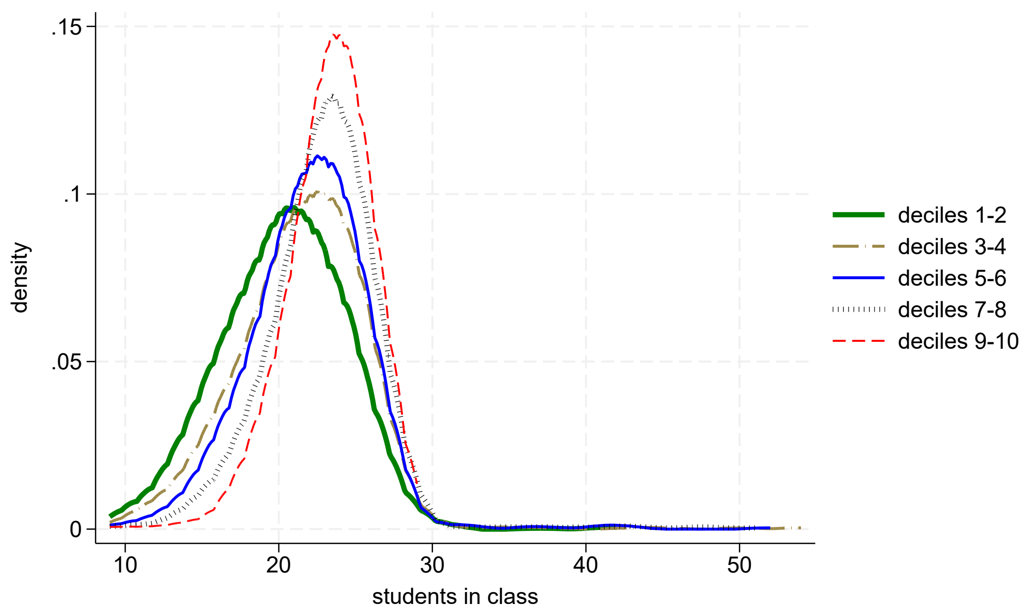
Figure 4: Quantile estimates by Gender



Note: Dependent variable is combined standardized test score in Danish in grade 6. 95% confidence intervals. Based on the with-in school between-class model.

Next, the sample is split by class size. Like teacher sorting, class size also reflects school quality. Most students have more than one school within relatively low distance and the most popular schools therefore have more students in each class. Since public funding follows the student this means that each schools financial situation can be improved upon when class sizes increase.[23] Schools where the average parental income is high have higher class size. The pattern is clear and monotonous in school SES deciles, see Figure 5.[24]

Figure 5: Average class size, by school SES decile



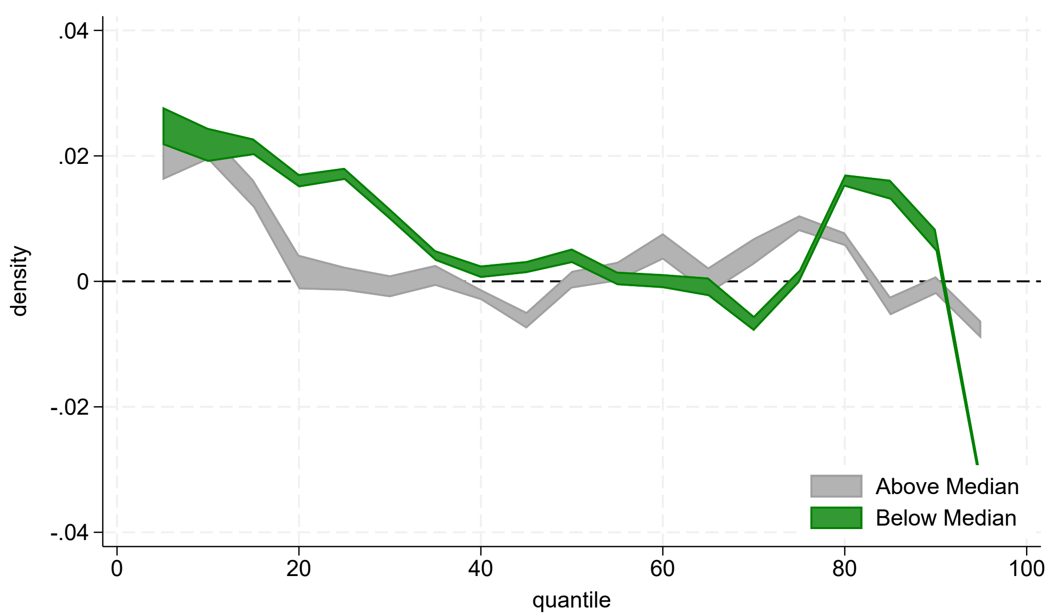
In order to empirically investigate the importance of class size for teacher effects, I split the sample at the median of the class size to study any difference between large and small class sizes. The parameters are again identified from within school between class variation. The results from two separate estimations of college-teacher effect on student outcome – for class size above and below the median respectively – are shown in Figure

[23]Students with special circumstances, including severe social problems, may come with more public financed resources but will also require extra resources. Exactly how many public resources there are allocated to each student is unobserved in the data.

[24]In other countries, class size is often lower in schools situated in affluent areas. In Denmark, each house belongs to a school catchment area, but if the maximum class size is not reached by students from the catchment area, students (or their parents) from other areas are offered a free school choice, so the most attractive schools tend to have larger class sizes.

6. The green confidence band represents small class sizes, which are also largely low SES schools, and the top 50% class sizes are represented by the grey band in Figure 6. College-educated teachers have a positive effect at the bottom of the student test score distribution irrespective of class size. For class sizes above the median, the positive effect of college-educated teacher is only found at the bottom 20% while the college-teacher impact becomes positive for the bottom 40% benefit when class sizes are below the median. These results appear reasonable. A college educated teachers are equipped with didactic skills that enables them to instruct low-performing students. This finding suggests that the positive impact a college-educated teacher may have (vis-à-vis a non-college educated teacher) also depends on the class size. The teachers attention-span can include a larger proportion of students when the class size is low.

Figure 6: Quantile parameter estimates of college-teacher on student outcome, by class size



Note: Dependent variable is combined standardized test score in Danish in grade 6. 95% confidence intervals. Based on the with-in school between-class model and Maimonides rule.

6 Discussion and conclusion

There is a worldwide shortage of qualified school teachers working as teachers in elementary school, and in many countries this problem has been rising over the past many years. This underscores the importance of understanding the prevalence and consequences of non-college educated teachers on student outcomes and how teacher sorting, whereby the most-qualified teachers tend to teach high-achieving students, cause impediments for students struggling with dyslexia, especially if they attend schools in low SES areas.

New to the literature, this paper emphasizes the importance of heterogeneity among non-educated school teachers. To date, a large literature has documented how the best-qualified among college-educated school teachers teach high-ability students. Important for estimation of teacher quality impact on student test outcomes is how we treat the reference group of non-educated teachers. I find that non-educated teachers also are sorted. The high school GPA is much higher among non-college-educated teachers in high SES schools compared to low SES schools. Models that rely on within school identification therefore needs to allow for non-linearity in order to take this into account.

The impact of having a college educated teacher is positive and significant on average, albeit with a low parameter estimate. Across school SES levels very large differences are found and mostly so between the top 25% most affluent schools, measured by average parental income, and the remaining 75%. These results only corroborate existing findings such as Lankford et al. (2002), and Boyd et al. (2013) who find that low-income, low-achieving, and nonwhite students, particularly those in urban areas, often are taught by the least skilled teachers.[25]

In addition to these findings, that mirror previous findings based primarily on US

[25] A factor that likely contributes to the substantial gaps in academic achievement based on student income and ethnicity.

data, we here see how the positive impact of teacher credentials arises at the bottom 20-30% of the student test score distribution, and mostly in the low SES schools and/or the schools with low class size. Both aspects may in part reflect smaller number of students and thus class rooms where teacher skills better come to their right and can make more of a difference at the bottom of the test score distribution.

Students who are at-risk of dyslexia based on 4^{th} grade decoding score (in reading) generally progress less when tested in 6^{th} grade reading. Except for students from the top 25% school SES who make above-average progress between 4^{th} grade and 6^{th} grade. This is likely a reflection of both stronger emphasis on special instruction time for dyslexic students at the school combined with more affluent parents ability to assist their child and ability to buy private tutors. These likely explanations would benefit from data-based analysis. A natural suggestion for future research.

References

Aaronson, Daniel, Lisa Barrow, and William Sander. 2007. "Teachers and Student Achievement in the Chicago Public High Schools." *Journal of Labor Economics* 25 (1): 95–135. 10.1086/508733. 1, 12, 4.1

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge. 2023. "When should you adjust standard errors for clustering?" *The Quarterly Journal of Economics* 138 (1): 1–35. 27

AE Rådet, Copenhagen. 2023. "Næsten hver femte lærer i folkeskolen har ikke en læreruddannelse." 1

Andersen, Simon Calmar, Henrik Hansen, Lóa Björk Jóelsdottir, and Helena Skyt Nielsen. 2024. "The male shortcut to math achievement." *Available at SSRN 4766059.* 17

Beuchert, Louise, and Anne Nandrup. 2017. "The Danish National Tests at a Glance." *SSRN Electronic Journal.* 10.2139/ssrn.2996431. 9

Björklund, Anders, and Kjell G Salvanes. 2011. "Education and family background: Mechanisms and policies." In *Handbook of the Economics of Education*, Volume 3. 201–247, Elsevier. 1

Black, Sandra E, and Paul J Devereux. 2011. "Recent developments in intergenerational mobility." *Handbook of labor economics* 4 1487–1541. 1

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes. 2005. "Why the apple doesnt fall far: Understanding intergenerational transmission of human capital." *American economic review* 95 (1): 437–449. 1

Bonhomme, Stephane, Gregory Jolivet, and Edwin Leuven. 2016. "School Characteristics and Teacher Turnover: Assessing the Role of Preferences and Opportunities." *The Economic Journal* 126 (594): 1342–1371. 10.1111/ecoj.12279. 1

Boyd, Donald, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2005. "Explaining the Short Careers of High-Achieving Teachers in Schools with Low-Performing Students." *American Economic Review* 95 (2): 166–171. 10.1257/000282805774669628. 1

Boyd, Donald, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2013. "Analyzing the determinants of the matching of public school teachers to jobs: Disentangling the preferences of teachers and employers." *Journal of Labor Economics* 31 (1): 83–117. 6

Chetty, Raj, John N Friedman, and Jonah E Rockoff. 2014. "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates." *American economic review* 104 (9): 2593–2632. 12

Clotfelter, Charles T, Helen F Ladd, and Jacob L Vigdor. 2002. "Segregation and resegregation in North Carolina's public school classrooms." *NCL Rev.* 81 1463. 18

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2010. "Teacher Credentials and Student Achievement in High School: A Cross-Subject Analysis with Student Fixed Effects." *The Journal of Human Resources* 45 (3): 655–681, `http://www.jstor.org/stable/25703472`. 1, 26

Conger, Dylan. 2005. "Within-school segregation in an urban school district." *Educational Evaluation and Policy Analysis* 27 (3): 225–244. 18

Dee, Thomas S. 2007. "Teachers and the Gender Gaps in Student Achievement." *The Journal of Human Resources* 42 (3): 528–554, `http://www.jstor.org/stable/40057317`. 26

Elbro, Carsten, and Nicolai Kristensen. 2025. "Who is tested for dyslexia? Biases arising from social factors and expectations." *submitted.* 2.1, 3, 15, 20

Falk, Armin, Fabian Kosse, Pia Pinger, Hannah Schildberg-Hörisch, and Thomas Deckers. 2021. "Socioeconomic status and inequalities in childrenâs IQ and economic preferences." *Journal of Political Economy* 129 (9): 2504–2545. 1

Firpo, Sergio, Nicole M Fortin, and Thomas Lemieux. 2009. "Unconditional quantile regressions." *Econometrica* 77 (3): 953–973. 4.2

Gellert, Anna Steenberg. 2009. "Identifikation og undervisning af ordblinde med dansk som andetsprog." *København: Undervisningsministeriet.* 15

Gensowski, Miriam, Rasmus Landersø, Philip Dale, Anders Højen, Laura Justice, and Dorthe Bleses. 2024. "Public and parental investments, and children s skill formation." *Journal of Human Resources.* 5

Guarino, Cassandra M, Mark D Reckase, and Jeffrey M Wooldridge. 2015. "Can value-added measures of teacher performance be trusted?" *Education Finance and Policy* 10 (1): 117–156. 2.2

Guryan, Jonathan, Erik Hurst, and Melissa Kearney. 2008. "Parental education and parental time with children." *Journal of Economic perspectives* 22 (3): 23–46. 1

Hanushek, Eric. 1971. "Teacher characteristics and gains in student achievement: Estimation using micro data." *The American Economic Review* 61 (2): 280–288. 12

Hanushek, Eric A., and Steven G. Rivkin. 2010. "The Quality and Distribution of Teachers under the No Child Left Behind Act." *The Journal of Economic Perspectives* 24 (3): 133–150, `http://www.jstor.org/stable/20799159`. 1

Jackson, C. Kirabo. 2009. "Student Demographics, Teacher Sorting, and Teacher Quality: Evidence from the End of School Desegregation." *Journal of Labor Economics* 27 (2): 213–256. 10.1086/599334. 1

Kalil, Ariel, Rebecca Ryan, and Michael Corey. 2012. "Diverging Destinies: Maternal Education and the Developmental Gradient in Time With Children." *Demography* 49 (4): 1361–1383. 10.1007/s13524-012-0129-5. 1

Kalogrides, Demetra, and Susanna Loeb. 2013. "Different Teachers, Different Peers: The Magnitude of Student Sorting Within Schools." *Educational Researcher* 42 (6): 304–316. 10.3102/0013189X13495087. 18

Killewald, Alexandra, and Jonathan Bearak. 2014. "Is the Motherhood Penalty Larger for Low-Wage Women? A Comment on Quantile Regression." *American Sociological Review* 79 (2): 350–357. 10.1177/0003122414524574. 4.2

Lankford, Hamilton, Susanna Loeb, and James Wyckoff. 2002. "Teacher sorting and the plight of urban schools: A descriptive analysis." *Educational evaluation and policy analysis* 24 (1): 37–62. 6

Morgan, Phillip R, and James M McPartland. 1981. "The Extent of Classroom Segregation within Desegregated Schools.." 18

Nguyen, Tuan D., Lam D. Pham, Michael Crouch, and Matthew G. Springer. 2020. "The correlates of teacher turnover: An updated and expanded Meta-analysis of the literature." *Educational Research Review* 31 100355. 10.1016/j.edurev.2020.100355. 1

OECD. 2020. "results (volume V): effective policies, successful schools, PISA." 1

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica* 73 (2): 417–458. 10.1111/j.1468-0262.2005.00584.x. 1, 12

Sibieta, Luke. 2020. "Teacher shortages in England: Analysis and pay options." *Education Policy Institute*. 1

Snowling, Margaret J., and Monica Melby-Lervåg. 2016. "Oral language deficits in familial dyslexia: A meta-analysis and review.." *Psychological Bulletin* 142 (5): 498–545. 10.1037/bul0000037. 4

Solheim, Oddny Judith, Julie Arntzen, and Njål Foldnes. 2024. "Norwegian classroom teachers' and specialized "resource" teachers' dyslexia knowledge." *Reading and Writing* 37 (10): 2619–2641. 1

Taylor, Jeanette, Alysia D Roehrig, B Soden Hensler, CM Connor, and Christopher Schatschneider. 2010. "Teacher quality moderates the genetic effects on early reading." *Science* 328 (5977): 512–514. 1

A Appendix A: Within class identification

In the second econometric approach, I compare the within-class difference across classes exposed to teachers with different teacher college GPA. The impact of GPA is captured by the interaction term with the indicator for being "At Risk". This approach, therefore, is more robust to, say, school-leader-induced sorting of students to teachers (or vice versa). If the best teachers have more demanding students in their class, and if there are spillovers between class peers, this method will also suffer from downward bias in the teacher quality estimate, but remember-still that we also condition on lagged outcomes which likely reduces any such bias.

This second model-type specification can be written as

$$\begin{aligned} TestScore_{itc} = \beta_0 + \beta_1 AtRisk4_i + \beta_2 AtRisk4_i \cdot GPA_c + \\ \gamma_c + \theta_t + X_i\beta_4 + \left(AtRisk4_i \cdot X_i\right)\beta_5 + \varepsilon_{itc}, \end{aligned} \tag{3}$$

where γ_c signifies a classroom (teacher) fixed effect. The parameter of interest in equation (3) is β_2, which gives the causal impact of teachers college GPA on the teachers ability to lift students at-risk. Given the complexity involved in learning how to read, at least for poor readers, we expect to find that teachers with higher GPA from teacher college have a positive and significant impact on at-risk students. And, especially at low SES schools. The GPA is normalized so that we interpret the parameter as the effect of a one SD increase in teacher-college GPA.[26]

[26] A third estimation alternative would be to use a within-student between-subject identification strategy as in Dee (2007) and Clotfelter et al. (2010). This identification strategy is not well suited to our problem at hand as it relies on teachers in different subjects and dyslexic students have an innate disadvantage in reading and spelling (Danish) while they may perform much better in math even though some reading is also required in math and other subjects. The difference in performance between subjects, therefore, may arise for a variety of reasons and not easily isolated to teacher qualifications.

Estimation of within class models is limited to classes where the teacher has a degree from teacher-college and where the GPA is observed, i.e. completion after 2004. The within class estimation approach is therefore included mainly as a robustness check.

A.1 Results using within class identification

Using with-in class identification, the parameter of interest is the interaction term between college GPA and an indicator for being at risk of dyslexia.

Qualitatively similar results emerge when estimating the with-in class model. Standard errors are here clustered at the class level, which results in standard errors that render the estimate of teacher quality insignificant.[27] Note though, that the parameter estimate for at-risk students in low SES (Q1) schools (row 1 column 2) is 2 percent of a SD while it is -2 percent for high SES (Q4) schools.

Based on the within-class identification strategy the heterogeneous impact of being at risk emerges again. In the top 25% most affluent schools, the indicator for being at risk of dyslexia in grade 4 has a positive impact on test score outcomes in grade 6, conditional on the many controls and in particular conditional on previous test scores. The sign is reversed in the remaining distribution and highly significant for low SES schools. This is consistent with the results found for the within school - between class identification strategy shown in Table 6. Possible explanations for these results are given above.

[27] These standard errors are likely conservative. Abadie et al. (2023) develop a new and improved framework for clustering. They show that when the number of clusters in the sample is a non-negligible fraction of the number of clusters in the population (the case here), conventional clustered standard errors can be severely inflated. Unfortunately, the theory is not yet developed to also allow for many control variables, and adapting, say, the Frisch-Waugh-Lovell (FWL)-theorem does not apply in this setting, because if the treatment variable is residualized, it will no longer be binary, which is required in Abadie et al. (2023).

Table 5: Within Class Estimates Conditional on School SES

	(1) All	(2) Q1	(3) Q2-3	(4) Q4
College GPA x AtRisk4	0.00456	0.0206	0.000414	-0.0264
	(0.0104)	(0.0166)	(0.0157)	(0.0256)
Test score grade 4	0.411***	0.393***	0.416***	0.427***
	(0.00484)	(0.0103)	(0.00683)	(0.00904)
Test score grade 2	0.198***	0.188***	0.199***	0.201***
	(0.00448)	(0.00940)	(0.00634)	(0.00863)
AtRisk4	-0.0120	-0.0630**	-0.0115	0.0668**
	(0.0112)	(0.0208)	(0.0159)	(0.0253)
Observations	80,721	19,375	39,740	21,606

Note: Dependent variable is combined standardized test score in Danish in grade 6. Based on the within class model. Standard errors clustered at the class level in parenthesis. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

B Appendix B: Additional tables

Variable Name	Description
Merit	Equals 1 if merit education completed. Zero otherwise.
Teacher55+	Equals 1 if aged 55 or higher. Zero otherwise.
Mother low educated	Mother has no qualifying degree.
Father low educated	Father has no qualifying degree.
Classsize	Number of students in class.
Class_add	Number of students in class with an ADD diagnose.
School size	Number of students at the school.
Year dummies	Reference year is 2016.
At-risk 2nd grade	Equals 1 if in bottom 10% Nationwide of reading test. Zero otherwise.
At-risk 4th grade	Equals 1 if in bottom 10% Nationwide of reading test. Zero otherwise.
Teacher college educated×At-Risk4	Interaction of the two variables.
Merit×Risk4	Interaction of the two variables.
Test Score grade 4	Reading score in grade 2.
Test Score grade 2	Reading score in grade 4.

Table 6: Reading (Danish) Teacher characteristics, by school SES decile

	SES Quartile		
	1	2	3
Teacher-college educated teachers			
Age	45.5	45.4	44.2
GPA from teacher-college	7.4	7.5	7.9
GPA from teacher-college missing	59.8%	59.2%	53.8%
Experience (years)	17.6	17.2	15.8
Years of education	17.0	17.0	17.0
Two Teachers in Danish class	14.7%	12.8%	10.1%
Merit teachers			
Age	50.1	48.7	47.9
GPA from high school	7.0	6.5	7.0
GPA from high school missing	34.6%	30.9%	31.4%
Experience (years)	10.9	11.2	12.0
Years of education	17.1	17.1	17.2
Not teacher-college educated teachers			
Age	45.2	43.3	38.0
GPA from high school	7.1	7.3	7.8
GPA from high school missing	27.4%	19.6%	10.4%
Experience (years)	19.4	15.6	10.9
Years of education	14.8	15.7	15.7

Note: Teachers as observed in 6^{th} grade. There are 4,587 different Reading (Danish) teachers (often measured repeatedly over the years). Total number of students included here is 231,160 over the school years 2014/2015 to 2018/2019 and across 1,179 public schools. School SES deciles are based on average parental income.