STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
5350 Master's thesis in economics
Academic year 2019-2020

Picking up the PISA?

Teacher certification and student achievement in Sweden

Anja Tuneld (23697)

Abstract. In an effort to raise the status of teachers and increase the quality and equity of Swedish education by ensuring competent teachers, a national and subject-specific teacher certification was introduced in 2011, despite scant evidence in favor of teacher certification. This thesis attempts to answer the question: What effect do certified teachers in Sweden have on ninth grade student achievement? By applying fixed effects models to school-level panel data spanning the academic years 2014/2015-2018/2019, this study finds that certified teachers have significant and positive effects on student achievement in mathematics, at least in terms of standardized national test results. For the two other subjects studied, namely English and Swedish, no such effects are found. Consequently, teacher certification could be an effective way of boosting student achievement in certain subjects, whereas in others, it risks discouraging non-certified teachers of high quality from entering or remaining within the teaching profession. However, this thesis faces several limitations, why I encourage future research to further explore the relationship between teacher certification and student achievement within the Swedish context.

Keywords: Teacher certification, teacher quality, student achievement, education policy

JEL: I20, I21, I28, H52

Supervisor: Abhijeet Singh
Date submitted: 18 May 2020
Date examined: 26 May 2020
Examiner: Magnus Johannesson
Discussant: Martina Kaplanová

Acknowledgements

I would like to thank my supervisor Abhijeet Singh for his valuable feedback and guidance. I would also like to extend my gratitude to Felix Schafmeister for his helpful input.

Moreover, I would like to acknowledge the individuals at the Swedish National Agency for Education (Skolverket) who took their time to answer questions and provide me with additional data.

Further, I want to express my deep appreciation to the principals and teachers who were willing to share their knowledge and thoughts, while continuing their mission to educate the young.

Finally, I thank family and friends for their endless support and encouragement, not only during this thesis process, but throughout my studies.

Contents

List of Figures

List of Tables

1 Introduction

By the time the fourth wave of the Programme for International Student Assessment (PISA) surveys was published, it had become apparent that the worsening results of Swedish students was a trend. From the first wave in 2000, up until the fourth wave in 2009, the average score of Swedish 15-year-olds in this international standardized test had seen a continuous decline in reading, mathematics, and science. In mathematics, the students now scored below the Organisation for Economic Co-operation and Development (OECD) average. Furthermore, student equity was threatened. From being a top performer in terms of low variation between top- and bottom-performers, together with the rest of the Nordic countries, Sweden had fallen below the OECD average in this respect as well. The decreased equity was observed at both the student-level and at the school-level. In particular, OECD pointed towards the performance gaps between the genders and between immigrants and natives (Organisation for Economic Co–operation and Development, 2010; Taube et al., 2010).

The Swedish government had proposed a series of reforms concerning compulsory and upper secondary education, already prior to the fourth wave of PISA surveys. For instance, it was announced in 2008 that the entire grading system was to be replaced and that students would be graded from the sixth grade of compulsory school, rather than from the eighth grade (Swedish Government, 2008). Reforms that could potentially increase student achievement and equity was widely supported by the public, who observed each wave of the PISA surveys with intensifying dismay.

At the same time, Sweden was facing a growing disinterest towards the teaching profession among the general population. Since the end of the millennium, significantly fewer people were applying to enter pedagogical university education (Statistics Sweden, 2019). The National Union of Teachers in Sweden had for a long time pushed for increased professionalism within the teacher profession. They wanted to see subject-specific certification requirements in all schools, as a way to reward those who had obtained a formal teaching degree.

In October 2010, the government handed in the bill that would create a national teacher certification in Sweden[1]. The reform was executed stepwise, starting in July 2011 and ending in July 2015. From July 2011, individuals could start to apply for the certification and by July 2015, non-certified teachers could no longer get permanent employment or independently grade students. The Swedish National Agency for Education[2] administers the certification process in exchange for a fee of SEK 1,500 per application, paid by the applicant. Anyone with a teaching degree obtained in Sweden, will be granted certification in the subjects that their degree specifies (Swedish Government, 2010). For someone studying to become a teacher to children in the last three years of compulsory school (grade 7-9), this corresponds to at least 4.5 years of study to obtain certification within 2-3 subjects. In my sample, the mean share of certified teachers in the subjects mathematics, English, and Swedish across the schools was 75 percent by the end-line of the reform (in academic year 2015/2016).

The explicit objective of the reform was to raise the status of teachers and increase the quality and equity of Swedish education by ensuring competent teachers (Swedish Government, 2010). In practice, it sends a strong signal about who is deemed to be a high-quality teacher and discourages non-certified individuals from entering the profession. However, the evidence in favor of certified teachers is scant. In the US, where teacher certification and degrees are extensively researched, only mathematics teachers with subject-specific knowledge are found to boost student achievement (see e.g. Coenen et al., 2018; Wayne and Youngs, 2003). In Sweden, Andersson and Waldenström (2007) conclude that ninth grade students (i.e. students in their final year of compulsory school)

[1]"Lärarlegitimation" in Swedish.
[2]"Skolverket" in Swedish.

with a higher share of formally educated teachers[3] see an increase in their grade point average (GPA). Despite the substantial costs of introducing teacher certification, born mainly by individual teachers, schools, and municipalities, the effect of teacher certification has not been the subject of any study in Sweden. Hence, the research question that the rest of this thesis will be devoted to answer is: *What effect do certified teachers in Sweden have on ninth grade student achievement?*

By applying fixed effects models to school-level data about the core subjects mathematics, English, and Swedish, I find significantly positive effects of certified teachers on student achievement, but only within mathematics. The estimated effect of a school going from none to all of their mathematics teachers being certified, is an increase in average national tests results by 1.476 points, on a scale of 0-20 points. In English, I find no significant effects of certified teachers, whereas in Swedish, I find negative effects of certified teachers on average student achievement. The latter result is likely due to that there are two tracks within Swedish ("Swedish" and "Swedish as second language") and that certified teachers deem fewer students to be in need of the lower level track. Hence, schools with more certified Swedish teachers are more likely to have low-performing students participate in the higher-level track (which is included in this study). Consequently, I am unable to say anything about the effects of certified Swedish teachers on individual student achievement.

The remainder of this thesis will be organized as follows: Section 2 provides background information about the Swedish schooling system and the teacher certification. Section 3 reviews the previous literature, Section 4 discusses the data, and Section 5 describes the empirical approach. In Section 6, the main results are presented, which are followed up with further analysis in Section 7. Section 8 tests the robustness of the main results, Section 9 discusses the internal and external validity of the thesis, before Section 10 concludes.

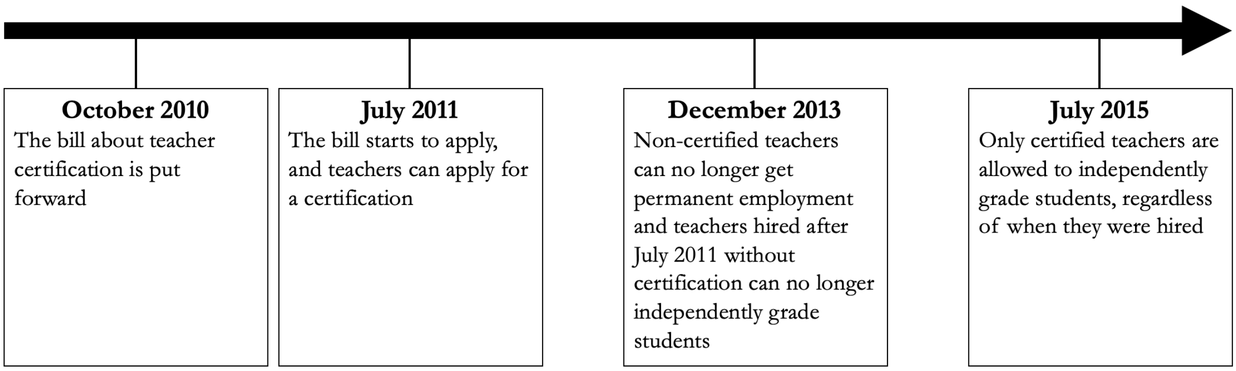
[3]Teachers with a formal teaching degree.

2 Background

Sweden has ten years of compulsory schooling, starting with the preschool class the year the child turns six, and ending with the ninth grade the year the student turns sixteen. The academic year starts in the fall and consists of two semesters. Optional upper secondary schooling consists of three years in either a vocational program or in a higher education preparatory program. All schooling is free of charge, regardless of whether the student attends a public or private school.

On October 7th 2010, the government handed in the bill that created the teacher certification in Sweden (Swedish Government, 2010). To give the schools enough time to prepare for the change, the policy was introduced stepwise. See Figure 1 for an overview. The end date of 2015 was announced already in the initial bill. Hence, schools should have been able to take steps towards ensuring that they would have enough certified teachers, already by the end of 2010.

Figure 1: Timeline of the introduction of teacher certification



Source: Author's rendering of information from the Swedish government (2010).

2.1 Teacher certification

From July 2011 and onward, teachers have been able to apply for a certification. To become certified, teachers need to send their application to the Swedish National Agency for Education at a cost of SEK 1,500. If the teacher is eligible for certification, it will specify both the subjects and the grades that the teacher gets certified within. There are two ways to obtain a certification: Either the teacher has a formal teaching degree obtained in Sweden, or the teacher has a non-Swedish teaching degree and have enough working experience or made other supplementary efforts that are equivalent to a Swedish teaching degree. The subjects and grades that the teacher gets certified within corresponds to his or her degree. For those studying to become teachers in grade 7-9, this is usually two or three subjects. The certification is automatically valid for lower grades, e.g. a teacher with certification for upper secondary school automatically becomes certified for grade 7-9 in compulsory school. Any teacher can get certified within more subjects if he or she has at least 8 years of experience in teaching the subject in question or has taken additional courses (e.g. at least 45 ECTS for mathematics). An effort to grant current teachers certification within more subjects has paralleled the reform by offering additional courses and granting funds to schools that have their teachers participate[4].

Since December 2013, schools are no longer allowed to permanently employ non-certified teachers. However, they can still get hired for one year at a time, given that there are no certified teachers within the organization that can do the job or any certified teachers applying for the position. Moreover, schools can end a teacher employment if the teacher is not certified by stating

[4]This effort is called "Lärarlyftet" in Swedish.

redundancy. Since July 2015, teachers without certification in any subject have to grade their students in collaboration with a certified teacher. This rule started to apply already in 2013 for teachers hired after 2011. If the certified teacher is certified within the subject in question, his or her opinion outweighs that of the non-certified teacher. In the case none of the teachers have certification within the subject, the principal should make the final decision (if there is a disagreement between the two teachers).

2.2 The Swedish schooling system

Most schools in Sweden are public and thus run by the municipalities (Båvner et al., 2011). Parents[5] have the option to apply to any school in the municipality but if they do not make an active choice, the child is automatically enrolled in the closest school that has available spots. If the parents wish for their child to attend a private school, they need to sign up to the specific school they want to attend. By the Eduation Act[6], all private schools (with some exceptions) should be open for enrollment to all children and the admission should be based on the principle of first come, first served. Hence, parents with foresight and knowledge about these rules often sign up their child years in advance. Since private schools are required to be free of charge, they are financed by public funds. However, they are allowed to be profit-driven, as opposed to public schools (Sahlgren, 2016).

Since 2011, students are graded each semester from sixth grade and onward on a scale of A to F with F being a failing grade and E the lowest passing grade. When these grades are aggregated, each grade is transformed into numerical value on a scale of 0 to 20 (see Table 1). From the time the students receive their first grades until graduation from compulsory school, they are graded in at least 16 mandatory subjects. However, many students graduate with 17 or 18 subjects, depending on if they choose to study a third language and/or study their mother tongue (granted to students with at least one parent speaking another native language than Swedish). In those cases, the subjects counting towards their total grade points are the 17 subjects that they perform the best in. For those taking only 16 subjects, those 16 subjects count towards their total grade points, resulting in a grade point ranging from 0 to 320 points. However, the grade point of those students taking more than 16 subjects, range from 0 to 340[7]. Only the grades of the final semester in ninth grade are included in the applications to upper secondary schooling. Thus, grades prior to the final semester serve the sole purpose of evaluation.

Table 1: National standards for converting alphabetical grades into numerical grades

Alphabetical grade	Numerical grade
A	20 points
B	17.5 points
C	15 points
D	12.5 points
E	10 points
F	0 points

Source: Author's own rendering of information from the Swedish Government (2008).

[5]For convenience, parent will refer to parent or any other legal guardian.

[6]"Skollagen" in Swedish.

[7]For the sake of comparison, I will be using the grade points average (GPA) throughout this study, by dividing the average total grade points of each school by 16.

To ensure relatively consistent grading across the schools, all schools are required to provide their students with the national tests that are administered by the Swedish National Agency for Education. These standardized, national tests are taken in third, sixth, and ninth grade of compulsory school. In ninth grade they include the subjects mathematics, English, and Swedish for all students. Moreover, each school is randomly allocated one of the natural science-subjects (biology, chemistry, or physics) and one of the social science-subjects (geography, religion, history, or civics). The tests consist of several (typically both written and oral) parts that are taken on different dates. The teachers have no access to the tests before the test dates but do correct them, using a thorough manual on how to assign points to each question. To check that teachers actually correct the tests according to the manual, a random sample of tests are sent to the Swedish National Agency for Education for a second round of grading. When all parts of each test have been corrected, the points are converted into a single (alphabetical) grade. The result of the national test is supposed to serve as a guideline for teachers when grading the students, but they are not bound to grant the student the same grade as the one received on the national test (Swedish National Agency for Education, 2020a).

3 Literature review

This section summarizes and discusses previous literature regarding teacher quality, focusing especially on the effect of teacher certification and degrees on student achievement.

3.1 Determinants of student achievement

The question of what factors determine student achievement, or go into the educational production function, has long caught the attention of economists, sociologists, and policy makers. In simple regression frameworks, much of the variation in student test scores is explained by individual and family background traits. Although other variables such as per student spending, class size, and teacher quality are of higher relevance to policy makers, conclusive results regarding their effect on student achievement have been scarcer. In a 1986 review of the studies on educational productivity at the time, it was concluded that: "differences in [school] quality do not seem to reflect variations in expenditures, class sizes, or other commonly measured attributes of schools and teachers" (Hanushek, 1986, p. 1142).

However, this puzzling conclusion did not discourage from continued study of the relationship between different schooling inputs and educational outcomes. Later authors have to a greater extent been able to take advantage of experimental, quasi-experimental, and econometric methods such as fixed effect estimations to reach more convincing results. An example is given by Angrist and Lavy (1999) who use a regression discontinuity design to study the effect of class sizes, finding positive effects of smaller classes on student test scores in mathematics and reading. Another example is given by Krueger (1999) who uses data from the STAR-experiment where children were randomly assigned to different types of classes. Just as Angrist and Lavy, he finds that smaller classes improve test scores. Moreover, he finds statistically significant, albeit small, positive effects of teacher experience. This finding is partly confirmed by Rivkin et al. (2005) who only find positive effects of teacher experience in the initial years of teaching. Many authors have studied other aspects of teacher quality such as teaching degrees, certification, and other formal requirements. The rest of this section will focus specifically on those other aspects of teacher quality.

3.2 The challenges associated with the study of teacher quality

The most apparent challenge facing all studies using non-experimental data is the potential non-random sorting of teachers. This can occur both across regions, schools, and classes within schools. For instance, Ingersoll and Gruber (1996) and Wayne (2002) have shown that low-income students in the US have fewer teachers with certain characteristics. A failure to control for this non-random sorting would lead to an upward (downward) bias if high-quality teachers teach high- (low-) performing students to a larger extent.

A second problem, namely multicollinearity, can arise when researchers try to control for variables such as student characteristics. This problem is a result of the typically high correlation between various student characteristics and some measures of teacher quality (Dewey et al., 2000). Multicollinearity makes the coefficients imprecise and often results in a failure to detect any statistically significant relationship between teacher quality and student outcomes.

One, in theory simple, solution is to study the relationship between teacher quality and student outcomes through experiments. By randomly assigning students to teachers (or vice versa), one can ensure that the sorting of teachers is random. Moreover, there is no need to control for variables that might lead to problems with multicollinearity (given that the randomization process has truly been random). However, running these types of experiments is all but simple in practice. Not only will anyone trying to conduct an experiment face ethical and economical constraints, but find difficulties in making the participants (i.e. schools, teachers, students, and parents) comply to the rules of the experiment. Hence, only a handful of experimental studies on teacher quality exists (see e.g. Cantrell et al.; Glazerman et al., 2006; Krueger, 1999). Even when an experiment is carried out, the random allocation is rarely done perfectly. For instance, Krueger (1999) uses statistical methods to account for the transitions between classes and the nonrandom attrition that occurred during the STAR-experiment.

Lastly, the majority of studies on teacher quality are carried out in the United States. The results from these studies are unlikely to be transferable to other countries with other educational settings. Additionally, rules and legislation differ across states. For instance, the rules outlining what constitutes a certified teacher differ across states. Thus, any study using national data on state-certified teachers is bundling different things into one category.

3.3 The National Board for Professional Teaching Standards

Given the lack of harmonization of teacher certification across the US states, several studies have focused on teachers certified by The National Board for Professional Teaching Standards (NBPTS). The requirements to become an NBPTS teacher is the same across the US and many states provide higher salaries to NBPTS teachers. To apply for the certification, at least three years of teaching experience is required. The application process includes a written component, recorded videos, and a six-part content knowledge assessment.

Most studies on the effect of NBPTS teachers apply varying models on panel data of students and teachers at the individual level. Cantrell et al. (2008) use an experimental approach and find positive but imprecisely estimated effects of NBPTS teachers on student performance. Others use observational data to run value added models including school-fixed effects (Anthony and Goldhaber, 2007; Harris and Sass, 2009), student-fixed effects (Clotfelter et al., 2007), or controls for lagged student achievement (Cowan and Goldhaber, 2016). Moreover, Horoi and Bhai (2018) add family and neighborhood shocks by including twin- and sibling-fixed effects.

All the referenced studies examine the effect of NBPTS teachers on student achievement in mathematics and reading (at least). Some find none or very small effects, such as in the case of Harris and Sass (2009), who only find positive effects for certain cohorts, tests, or grade levels. Others find that having an NBPTS teacher leads to an increase of up to 0.05 of a standard deviation of the standardized test-score gains distribution. Although there is no consensus regarding the general effects of NBPTS teachers on student achievement, the studies that have investigated the mechanisms behind any effect, overwhelmingly support the signaling hypothesis rather than the human capital hypothesis[8] (e.g. Anthony and Goldhaber, 2007; Chingos and Peterson, 2011; Clotfelter et al., 2007; Harris and Sass, 2009).

3.4 Degrees and other teaching certifications

The tertiary education of teachers and certifications more directly linked to their educational background have also been the subject of several studies, mainly in the US. One early study (Hawk et al., 1985) found that having mathematics teachers with subject specific certification in mathematics lead to higher student achievement. However, it failed to control for the socioeconomic background of the students. Luckily, later studies have been able to overcome this shortcoming.

Given that certification requirements vary across states, it is common for studies to take place within single states or cities. As an example, Kane et al. (2008) study fourth to eighth graders

[8] The signaling hypothesis in this case states that the certification process leads to only high-quality teachers applying. That is, the certification is difficult or tedious enough to signal that teachers that have gone through it is of high-quality. The human capital hypothesis, on the other hand, states that the certification process actually makes the applicants better teachers, by improving their human capital. That is, important teaching skills are acquired by the teachers during the process of getting certified.

in New York City, finding no statistically significant effect on student achievement of neither academic credentials of teachers nor teacher certification. One pair of authors who has managed to study degrees and certifications with national data, is Goldhaber and Brewer (1997a, 1997b, 2000). They have studied the impact of both degrees and teaching certifications on tenth and twelfth grade achievement and found that students in mathematics learn more from teachers with subject-specific degrees or subject-specific certifications. However, this relationship between subject-specific knowledge and student test scores was not found in the other subjects examined (English, science, and history). Additionally, Monk and King (1994) also use national data to study the relationship between courses taken by teachers and student achievement and find positive correlations within mathematics.

To summarize, several authors have found a positive relationship between mathematics knowledge of teachers and student achievement in mathematics. However, there is little evidence that this holds true within other subjects. This view is confirmed by two systematic reviews on the subject (Coenen et al., 2018; Wayne and Youngs, 2003). These reviews conclude that "teacher certification, in general, has no positive effect on student performance, but subject-specific certification, especially in math, is frequently found to be positively related to student performance" (Coenen et al., 2018, p. 873) and "in the case of degrees, coursework, and certification, findings have been inconclusive except in mathematics, where high school students clearly learn more from teachers with certification in mathematics, degrees related to mathematics, and coursework related to mathematics" (Wayne and Youngs, 2003, p. 107).

3.5 Evidence from Sweden

Few studies linking teaching degrees or certification to student achievement exist within the Swedish context. However, one such study links formal teaching degrees to third grade reading achievement, finding positive effects of having educated teachers in the early years. However, no significant effects of teacher experience is found (Myrberg, 2007). Another relevant study is is authored by Andersson and Waldenström (2007) and focuses on the relationship between teachers with a formal teaching degree[9] and the GPA of ninth graders. Using a model with school fixed effects and another with instrumental variables (constructed by teacher unemployment and a temporary grant), they conclude that an increase of educated teachers by one percentage point leads to an increase in percentile ranked GPA of 0.6 units. As opposed to the American studies, Andersson and Waldenström identify overall positive effects of having teachers with teaching degrees. However, it is important to note that their dependent variable is not based on standardized test scores but on grades which are subject to teacher bias.

[9]The authors refer to teachers with degrees as "certified", although a formal teaching certification did not exist at that point in time.

4 Data

This section describes the data and discusses the possible issues that attrition and measurement errors present.

4.1 Data description

In this study, I exploit the variation in the share of certified teachers within compulsory schools from academic year 2014/2015 to 2018/2019. I will refer to the year when the academic years start, i.e. the data set includes the years 2014-2018. The data set is compiled using public data made available by the Swedish National Agency for Education (2020b). It includes all schools that have valid data during 2014-2018 of the GPA of graduating ninth graders. This amounts to 1,039 schools and includes both schools that offer all grades (i.e. all years of compulsory schooling) and schools with more limited offerings (e.g. only grades 7-9). Summary statistics of relevant variables are presented in Table 2. It includes school-level variables concerning the entire school (or all subjects), variables that are subject-specific to each school, and variables that only concern the ninth graders at each school. Unless stated in the variable name, the variables concern all students at the school.

Table 2: Summary statistics

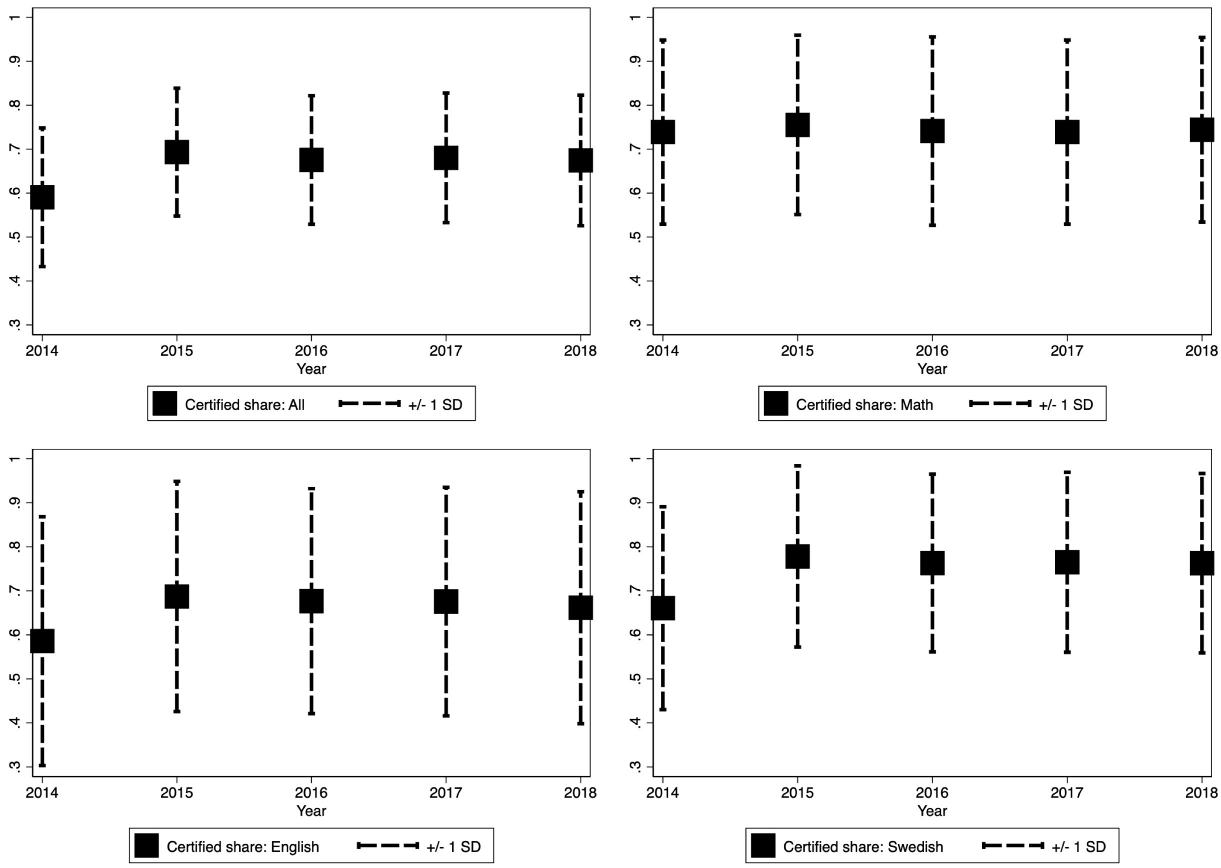
	All		Math		English		Swedish	
Variables	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Private ownership dummy	0.24	0.43						
Number of teachers	26.4	11.5	3.87	2.05	2.41	1.15	4.09	2.58
Share of certified teachers	0.68	0.14	0.75	0.20	0.67	0.25	0.76	0.20
Share with pedagogical degree	0.82	0.11						
Share of permanently employed	0.84	0.11						
Number of students	348	172						
Share of educated parents	0.55	0.17						
Share of immigrants	0.24	0.20						
9th grade: number of students	72.7	41.7	71.7	41.2	73.3	41.9	64.2	38.0
9th grade: share of females	0.48	0.076	0.48	0.076	0.48	0.076	0.49	0.077
9th grade: grade	14.2	1.74	12.2	1.85	14.4	1.94	14.2	1.40
9th grade females: grade	15.1	1.72	12.5	1.92	14.8	1.95	15.3	1.36
9th grade: passing grade share	0.77	0.15	0.90	0.091	0.92	0.078	0.96	0.041
9th grade females: passing share	0.81	0.13	0.91	0.090	0.93	0.077	0.97	0.037
9th grade: National test			11.1	2.21	15.0	1.58	13.7	1.46
9th grade females: National test			11.2	2.32	15.1	1.66	14.7	1.38
9th grade: Bias			1.18	1.26	-0.62	1.00	0.48	0.80

Notes: 1,039 schools are included over five years (2014-2018). All variables are recorded at the school-level. *All* refers to the entire school, i.e. all 16-18 subjects in compulsory school. When subject-specific data for mathematics, English, and Swedish are available, they are displayed. *Share of educated parents* is the share of students with at least one parent with tertiary education. *Share of immigrants* is the share of students not born in Sweden or without any parent born in Sweden. *Bias* is constructed as the difference between grade and national test results ($Grade - National\ test$).

In the overall sample, 24 percent of the schools are privately owned. On average, the schools in the sample has 26 teachers of which 84 percent are permanently employed, 82 percent has a pedagogical degree, and 68 percent are certified in at least one subject[10]. Naturally, more teachers have a degree compared to having a certification, since a degree is required to get certified. Considering the three core subjects (mathematics, English and Swedish), schools have fewer teachers in English than in mathematics or Swedish. This can be explained by the fact that fewer hours are devoted to English during compulsory school. Moreover, mathematics and Swedish teachers are certified within their subject to a greater extent than English teachers.

Turning to the students, there are on average 348 students at each school, although this is highly variable. Of these, 73 are in the ninth grade and their study achievements are recorded in this study. The number of students that were recorded in mathematics and English varies slightly whereas on average 8.5 students (around 12 percent) do not take Swedish. This is because students that are not native speakers of Swedish can take another subject termed "Swedish as second language", and are subsequently not graded in Swedish. The results in the subject "Swedish as second language" are not included in this sample. At the schools, on average 55 percent of the students have at least one parent with tertiary education and 24 percent of the students are not born in Sweden or have no parent born in Sweden.

Figure 2: Mean and standard deviation of the share of certified teachers per subject and year



Notes: *All* refers to teachers across the entire school, i.e. all 16-18 subjects in compulsory school. Each year refers to an academic year, e.g. 2014 refers to academic year 2014/2015. Source: Author's rendering of data from the Swedish National Agency for Education (2020b).

[10]The fact that more teachers are permanently employed than certified may seem odd. However, it should be noted that it was possible to gain permanent employment without certification before 2013.

When it comes to ninth grade student achievement, 77 percent of the students graduate from compulsory school with a passing grade in all 16-18 subjects. The average GPA is 14.2 points which corresponds to receiving C's and D's (with slightly more C's). Female students tend to have higher grades than their male counterparts. In mathematics, the students on average get lower grades compared to their GPA's but their mathematics grades are still higher than their performance on the national tests, where they only score slightly better than the passing grade (E). This leads to a large positive bias of 1.18 points (on a scale of 0 to 20) when grading and national test results are compared. However, the standard deviation is high, indicating that this is highly variable across schools and years. In English, the grades are in line with the GPA, but students outperform that on the national tests, leading to a negative bias. That is, although the students perform well on national tests, the teachers consistently grant them lower grades. Also in Swedish the grades are in line with the overall GPA. However, their national test results are slightly lower, resulting in a positive bias of 0.48 points.

Figure 2 displays four graphs showing the variation in the share of certified teachers across the years 2014 to 2018 (in the schools as a whole and in each core subject). All graphs except for mathematics show a clear increase in certified teachers between 2014 and 2015. It was also during this period that the reform was fully implemented, and all non-certified teachers lost their right to independently grade students. Based on the data available, it is not clear if this increase stems from an increase of qualified teachers, or if many teachers waited until the last minute to apply for their certification. This issue will be discussed further in Section 8.3. As for the years 2015-2018, there is no clear trend in the variation of certified teachers. This is notable since the teacher certification reform was meant to attract more certified teachers to the profession. However, it is possible that the share of certified teachers increased between 2011 and 2014.

4.2 Attrition and measurement errors

In the academic year 2018/2019, there were 1,664 schools in Sweden that taught ninth graders. As mentioned in the previous sub-section, only 1,039 of them are included in my sample. The reason for this attrition is that schools either are not found in the data for all five years (2014-2018) or have missing values of the grade point average of graduating ninth graders.

Of the schools operating in 2018, 357 of them are excluded from the sample since they are not recorded in the previous four years. Hence, those schools are newly opened and have been in operation for less than five years. The rest of the excluded schools, 272 in total, have missing values in at least one of the five years. The majority of these schools have missing values because they have less than ten students in the ninth grade. In those cases, the Swedish National Agency for Education protects the integrity of the students by not publishing the average grades. In each year between 2014 and 2018, between 87 and 94 percent of the missing values are due to that the schools have less than 10 students in ninth grade. The rest are missing for unknown reasons.

Attrition does not present any issues as long as it is random, which is the same as saying that the selected sample is random. In this study, the sample is not random since it selects larger schools that have been around for some time. The question then becomes: Is it likely that an analysis based on this non-random sample would yield biased estimates of the effect of teacher certification? It is possible that larger schools can allocate their certified teachers more efficiently across classrooms. Additionally, larger schools are situated in more urban areas where the student characteristics may vary from smaller schools. This will be tested for in Section 7.2-3, without finding that the exclusion of smaller schools would significantly change the estimates. When it comes to the age of the schools, I have no possibility to test whether certified teachers have differential effects depending on for how long the school has been operating.

Nevertheless, it is evident that the share of private schools is underrepresented in my data, as a result of attrition. Private schools tend to be both smaller and newer, compared to public schools. Of all compulsory schools teaching ninth graders in Sweden, the share of private schools has been increasing, from 29 percent in 2014 to 31 percent in 2018. In my sample, the average share of private schools is only 24 percent. In Section 7.1, I find that the effects of certified teachers are greater for private schools, indicating that attrition bias my results towards zero.

On a related note, national tests data in mathematics are missing for all schools in academic year 2017/2018. This is because there were indications that some parts of the national tests had been leaked in advance. As a result, schools that suspected that their students had knowledge of the leakage, were urged to use a replacement test. Since 76 percent of the schools used the replacement test, the Swedish National Agency for Education decided to not publish the results of any school.

When it comes to the accuracy of the data, it is never possible to completely guard against measurement errors. Misreporting and coding errors can, and typically do, occur. Measurement errors in the explained variable, that are independent of the explanatory variables, do not lead to any bias in the coefficients. Random measurement errors in the explanatory variables, on the other hand, will bias the coefficients towards zero, or no effect. Hence, the main worry concerns the measurement of the share of certified teachers. Each school is required to report to Statistics Sweden who their teachers are and what courses they are teaching. Incongruous records are double checked with the schools in question. Then, Statistics Sweden runs this information against the official register of teacher certifications, determining if the teachers are teaching in subjects they are certified within. Lastly, proper adjustments are made for those teachers not teaching full-time. Mistakes can be made during this process and in the end, the schools are responsible for providing Statistics Sweden with correct information. One source of error regards the timing of data collection. Schools report their data in October, stating what subjects the teachers are teaching and will teach in during the entire academic year. If any changes in subjects taught occur during the year, they will not be picked up in my data. Hence, it is possible that measurement errors in my main explanatory variables create a bias towards zero. That being said, all of my data is official data published by the Swedish National Agency for Education and collected by Statistics Sweden, ensuring that it is of relatively high quality.

5 Empirical approach

Student achievement can be measured in numerous ways. Rather than aggregating student achievement across subjects by using measures such as GPA, I have decided to focus solely on the three core subjects mathematics, English, and Swedish. Within these three subjects, I am able to not only look at the average grade of the ninth graders and the share of ninth graders passing each subject, but also the average score on standardized national tests. Within education research, standardized test scores have long been the preferred measure of student achievement, since it is less likely to be subject to teacher bias.

More than studying the three measures of student achievement mentioned (grade, share passing, and national test results), I will examine the effect of certified teachers on the difference between grade and national test, termed bias throughout this thesis. The rationale behind this, is that teachers are required to pay particular attention to the result on the national tests when granting grades to their students. Moreover, the national tests are constructed in such manner that they closely follow the grading criteria. Unless the teachers exhibit a positive bias towards their students or create exams and assignments that are at a lower difficulty level than stated by the grading criteria, the granted grades should be the same or lower than the national test results (given that some students have lower ambitions than their knowledge level). However, a positive bias was detected in the subjects mathematics and Swedish in my sample. Other authors have also noted and studied the inflationary pressure on grades in Sweden (see e.g. Wikström and Wikström, 2005). Hence, it is of interest to study if certified teachers make different grading decisions compared to their non-certified peers.

As mentioned in the literature review, the main issue facing researchers studying teacher quality and student achievement is the non-random sorting of teachers across schools. In the case of teacher certification, it is plausible that better organized schools with high-performing students hire more certified teachers. This could both be the result of higher performing schools being better at recruiting certified teachers, and that certified teachers seek out good schools to teach in. A second problem, specific to this study, is that the sample includes the academic year 2014/2015 when the teacher certification reform was not fully implemented. It is possible that better organized schools urged their teachers to already have their certification ready by this year, whereas there was an under reporting of certified teachers at other schools in this initial year.

A straightforward way to control for the differing characteristics of schools is to include school fixed effects. Hence, the following specification is proposed:

$$Y_{stj} = \alpha_0 + \alpha_1 Certified_{stj} + \delta_s + \delta_t + \beta \mathbf{X}_{stj} + \varepsilon_{stj} \tag{1}$$

where Y_{stj} is the outcome of interest at school s in time t in subject j. $Certified_{stj}$ is the share of certified teachers, δ_s refers to school fixed effects, δ_t refers to time fixed effects, $\mathbf{X}_{stj}$ is a vector of controls including time-varying student characteristics at the school-level, and ε_{stj} is the error term. The variables included in the vector of controls are the share of female students in ninth grade, the share of students not born in Sweden or with (both) parents not born in Sweden (at the school), the share of students with at least one parent with tertiary education (at the school), and the number of ninth graders taking the subject in question. Robust standard errors are presented, which correct for heteroskedasticity, including clustering at the school level.

This specification uses the variation in one subject within schools over time, i.e. each subject is studied separately. The coefficient of interest in this specification is α_1, which estimates the effect of certified teachers in the subject in question. Although the specification controls for fixed school effects, it is not possible to control for time-varying school effects without adding a vector of controls. Hence, the vector of controls needs to be rich enough for the interpretation of α_1 to be causal, or the specification will suffer from omitted variables bias. Any spillover effects of certified teachers between subjects would cause omitted variables bias since the share of certified teachers in other subjects are not included in the vector of controls. The assumption of no spillover effects of teachers across subjects means that an additional teacher in e.g. mathematics is allowed to have effects on mathematics achievement, but not on achievement in any other subject.

I deem this first specification to be highly susceptible to omitted variable biases, since my vector of controls only include four variables. Since it is reasonable that the school fixed effects are the same for every school, regardless of which core subject is studied, I will pool all three subjects into one data set. Hence, a second specification that exploits the variation across subjects within schools and within years is proposed:

$$Y_{stj} = \alpha_0 + \alpha_1 Certified_{stj} + \rho_j + \gamma_{st} + \varepsilon_{stj} \tag{2}$$

where Y_{stj} is the outcome of interest at school s in time t in subject j. The subjects included are mathematics, English, and Swedish and these are all run in a single regression. $Certified_{stj}$ is the share of certified teachers, ρ_j refers to subject fixed effects (i.e. one dummy for Swedish and one for English), γ_{st} refers to school-by-year fixed effects (i.e. there are five dummies per school), and ε_{stj} is the error term. The coefficient of interest, α_1, estimates the average effect of certified teachers in these three subjects. The identifying assumption is that there are no spillover effects of certified teachers between the subjects. If there are positive (negative) spillover effects, that would bias the estimates towards (away from) zero, since the share of certified teachers across the subjects correlate positively.

By including school-by-year fixed effects, I control for each school every year, meaning that both time-invariant and time-varying unobserved effects of schools are captured. As an example, these fixed effects would accurately capture the effect of a single school having a low-quality principal in a single year. The reason for including this type of discrete yearly controls, rather than including a continuous time trend for each school, is exactly to capture discrete events such as the replacement of a principal. The subject fixed effects control for different general levels across subjects. For instance, it was pointed out in Section 4.1 that ninth graders earn lower grades and score fewer points at the national test in mathematics compared to English and Swedish. This specification could still be subject to bias, if there are time-varying variables within the subjects that correlate with both the share of certified teachers and student achievement.

Specification 2 implicitly assumes that certified teachers within the core subjects (mathematics, English, and Swedish) all have the same effect. Seeing that previous literature has only identified certified mathematics teachers as significant contributors to student achievement, I will propose a second specification that allows for heterogeneous effects of certified teachers with respect to subjects:

$$\begin{aligned} Y_{stj} = \alpha_0 + \alpha_1 Certified_{stj} &+ \alpha_2 (Certified * English)_{stj} \\ &+ \alpha_3 (Certified * Swedish)_{stj} + \rho_j + \gamma_{st} + \varepsilon_{stj} \end{aligned} \tag{3}$$

where $English$ and $Swedish$ are indicator variables for respective subjects. This third specification is identical to Specification 2, except that two interaction terms have been added. Now, there are three coefficients of interest: α_1 alone should be interpreted as the effect of certified mathematics teachers on mathematics achievement, $(\alpha_1 + \alpha_2)$ should be interpreted as the effect of certified English teachers on English achievement, and $(\alpha_1 + \alpha_3)$ should be interpreted as the effect of certified Swedish teachers on Swedish achievement.

6 Main results

The results of running Specification 1 with national test results as outcome variable can be seen in Table 3 and the effect on the rest of the outcome variables can be seen in the Appendix, Tables A1-A3. The tables display the results both including and excluding the vector of controls. Since the coefficients of interest change when the controls are added, the analysis will be based on the results from the regressions including the vector of controls. Still, it is likely that there are unobserved variables that lead to biased coefficients. Consequently, I will treat the results of specification 1 as a benchmark, against which Specifications 2 and 3 can be compared. Using specification 1, I find no significant effect of teacher certification on grades or share of students passing in any subject (Tables A1 and A2). Looking at national test results on the other hand (Table 3), I find that certified teachers in mathematics have a positive effect, significant at the 1%-level. In English and Swedish, the effects of certified teachers are not statistically significant at any conventional level.

Table 3: Effect of teacher certification on national test results (0-20 points) per subject

Variables	(1) Math	(2) Math	(3) English	(4) English	(5) Swedish	(6) Swedish
Share of certified teachers	0.631***	0.446***	0.0850	0.114	-0.0994	-0.0532
	(0.172)	(0.168)	(0.0755)	(0.0799)	(0.118)	(0.135)
Observations	4,012	3,481	4,986	4,277	4,699	3,933
R-squared	0.248	0.286	0.012	0.029	0.029	0.060
Year FE	x	x	x	x	x	x
School FE	x	x	x	x	x	x
Controls		x		x		x

Data on national test results in mathematics are missing for 2017. The controls include the share of female students in ninth grade, the share of students not born in Sweden or with (both) parents not born in Sweden (at school-level), the share of students with at least one parent with tertiary education (at school-level), and the number of ninth graders taking the subject in question. Cluster-robust standard errors are in parenthesis. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

The results of running specifications 2 and 3 on the entire sample can be viewed in Table 4. In odd columns, Specification 2 is run, and in even columns, Specification 3 with the interaction terms is run. Across the core subjects, certified teachers have no statistically significant effects on grades (Column 1). However, when allowing for heterogeneous effects of certified teachers (Column 2), there are some interesting effects. A school going from no certified teachers to all certified teachers in mathematics, will in that same year see an increase in the average mathematics grade of their ninth graders of 0.378 points on a scale of 0 to 20, ceteris paribus. The effect is significantly different from zero on the 1%-level. It should be noted that this estimate is based on much smaller changes in the share of certified teachers. On average, a school sees a change of 15 percentage points between two consecutive years. My estimate implies that such an average change would lead to an increase in mathematics grades of 0.06 points, which is not substantial. In English, the effect of certified teachers is not statistically significant. The F-test testing the null hypothesis that $(\alpha_1 + \alpha_2) = 0$ has a p-value of 0.708. In Swedish, certified teachers lead to a decrease in average grade by 0.466 points, significantly different from zero at the 1%-level. Turning to the share of ninth graders passing the core subjects (Columns 3 and 4), the pattern is identical to the one observed for grades.

Table 4: Effect of teacher certification on student achievement in the core subjects (mathematics, English, and Swedish)

Variables	(1) Grade	(2) Grade	(3) Pass	(4) Pass	(5) National test	(6) National test	(7) Bias	(8) Bias
Share of certified teachers	-0.0144	0.378***	0.00105	0.0169***	0.254***	1.476***	-0.296***	-1.089***
	(0.0684)	(0.111)	(0.00300)	(0.00514)	(0.0762)	(0.158)	(0.0726)	(0.135)
Interaction: Certified * English		-0.412***		-0.0169***		-1.561***		1.092***
		(0.134)		(0.00567)		(0.176)		(0.152)
Interaction: Certified * Swedish		-0.844***		-0.0336***		-1.827***		1.029***
		(0.141)		(0.00687)		(0.188)		(0.159)
F-test (p-value): English		0.708		0.9912		0.3336		0.9963
F-test (p-value): Swedish		0.0000		0.0023		0.0024		0.5564
Observations	15,193	15,193	15,144	15,144	13,697	13,697	13,575	13,575
R-squared	0.619	0.621	0.307	0.310	0.774	0.779	0.483	0.489
School-by-year FE	x	x	x	x	x	x	x	x
Subject FE	x	x	x	x	x	x	x	x

Notes: Grade and National test vary between 0-20 points with 10 points corresponding to a passing grade. Bias is constructed as the difference between Grade and National test ($Grade - National\,test$). *F-test (p-value): English* is the p–value of the F-test testing the null hypothesis that the sum of the coefficients of the share of certified teachers and the interaction between certified and English equals zero ($(\alpha_1 + \alpha_2) = 0$). *F-test (p-value): Swedish* displays the equivalent value but for the subject Swedish ($(\alpha_1 + \alpha_3) = 0$). Data on national test results in mathematics are missing for 2017. Cluster-robust standard errors are in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

When it comes to national tests, certified teachers in the core subjects have a positive effect on average (Column 5). When the interaction terms are added (Column 6), it is evident that this effect is entirely driven by the large effect of certified mathematics teachers. My estimate suggests that having only certified mathematics teachers rather than none, increases the average national test result in math by 1.476 points, again on a scale of 0 to 20. Considering that the standard deviation of national test results across schools is 2.1 on average, this effect is quite large. Even with just an increase in the share of certified mathematics teachers by 15 percentage points (the average year-on-year change for schools), my estimate suggests that the average national test results would increase by a tenth of a standard deviation. There is no significant effect of certified English teachers whereas certified Swedish teachers actually have a significantly negative effect on average national test results of 0.351 points, when the share of certified teachers goes from zero to one. This is puzzling. Although teacher certification could be an imprecise predictor of teacher quality, it is difficult to imagine that having Swedish teachers with a formal degree would lead to significantly poorer student performance. One possible explanation is that certified and non-certified teachers make differential decisions when it comes to which students should take "Swedish" as subject and which should take "Swedish as second language". If that is the case, certified teachers could affect the types of student included in my outcome variable of interest, leading to problems with endogeneity. This point will be further expanded upon in Section 8.2.

The effect of certified teachers on bias, or the difference between grade and national test result, is negative when the interaction terms are not included in the regression (Column 7). But just as in the case of national test results, this is entirely driven by the mathematics teachers. This can be seen in Column 8, where the effect of certified teachers on bias is only significantly different from zero for mathematics teachers. Seeing that in my sample, mathematics is the subject with the largest positive bias of 1.2 points on average, the identified negative effect of 1.089 points is substantial. No bias or a slightly negative bias, would imply that students receive grades that more accurately reflect their knowledge, and it seems like certified teachers contribute towards that objective. However, it should be noted that some or much of the bias observed in my sample could be a result of measurement errors.

In summary, my estimates from specification 3 suggests that having certified teachers in mathematics significantly increase student achievement and has positive effects on grading from a fairness perspective[11]. While going from 0 to 1 in the share of certified teachers increase average national test results in mathematics by as much as 1.476 points, average grades only increase by 0.378 points. This, in a subject that sees teachers systematically grant students higher grades than their performance on standardized national tests. In English, certified teachers have no statistically significant effect on any outcome, whereas certified Swedish teachers significantly decrease average grade, the share of students passing the subject, and national test results. It is notable that mathematics is the only subject studied that has positive effects on student achievement, at the same time as it is the only subject where the share of certified teachers has not increased during the years studied (2014-2018). When using the more straightforward approach presented in Specification 1, the only significant result is that certified mathematics teacher have a positive effect on national test results. However, specification 1 is likely to be subject to omitted variables bias due to unobserved time-varying school effects. Consequently, Specification 3 is my preferred specification. It should be noted that my results rely on multiple hypothesis testing, which carries an increased risk of false positives.

As mentioned in the previous section, any violation of the assumption of no spillover effects across subjects would bias the estimates. This bias could go in both directions since it is possible to imagine both positive and negative spillovers of high-quality teachers. As an example of positive spillovers, having a good teacher in mathematics can lead to more enthusiasm towards school as

[11] "Fairness" in this regard, refers to how accurately a grade reflects a student's knowledge level.

a whole, increasing student achievement in all subjects. On the other hand, one can also imagine that having a good teacher in mathematics, shifts the student's focus towards mathematics and away from other subjects, creating negative spillovers. Spillover effects will be further discussed and tested for in section 8.1.

The mechanisms behind these results cannot be deciphered through this analysis. However, previous literature has shown that subject-specific knowledge of the teacher contributes to enhanced student achievement, at least within mathematics (see e.g. Coenen et al., 2018 or Wayne and Youngs, 2003). In this case, teacher certification corresponds to subject specific knowledge. This is a result of that only mathematics teachers certified in mathematics, English teachers certified in English, and Swedish teachers certified in Swedish, are labeled as certified. Hence, it is likely that the certification process (i.e. getting a pedagogical degree in the subjects you are teaching) is increasing the human capital of teachers. But this is not to say that the certification could not also act as a strong signal of innate abilities and ambition. The fact that certified mathematics teachers seem to grade their students more fairly than their non-certified counterparts, could also be a result of their pedagogical degree, which has taught them how the interpret the grading criteria in a manner that is more aligned to the grading of national tests. But again, it could also be a question of signaling an innate sense of impartiality.

Another question that should be asked in the light of these results is: What differentiates the subject mathematics from English and Swedish within the school context? One possible answer is that mathematics is fundamentally different from the studies of languages, to such extent that the quality of the teacher matters more to student achievement. In ninth grade classrooms, the mathematics subject is mostly taught in a de-contextualized and abstract setting. This type of abstract thinking is seldom practiced outside the classroom. Consequently, the time spent in mathematics class, shapes the student's performance to a high degree. In language studies, on the other hand, external factors such as the language spoken at home and how much and what the student reads in his or her spare time, can greatly influence student achievement. Another possible interpretation is that skills and knowledge not taught or emphasized during formal teaching training are crucial to student learning in language-subjects. If those skills are not taught and individuals pursuing teaching degrees are not more likely to exhibit these skills as part of their innate abilities, non–certified teachers could be of equal quality.

7 Further analysis

In this section, further analysis of the effect of teacher certification is undertaken by dividing the data set into sub-samples that are analyzed individually. The objective is to distinguish whether certified teachers have heterogeneous effects depending on what type of students or schools they teach in.

7.1 Private and public schools

Differential outcomes between private and public schools have long been of interest to scholars, and Swedish scholars have been no exception. However, the context in which private schools operate is quite different in Sweden compared to other countries. As described in Section 2.2, private schools are publicly funded and generally open to all students. Given that private schools get paid per student (with school vouchers covering all students), there is a general fear that private schools are more generous in their grading, in an effort to please current and future parents. Indeed, at the upper secondary schooling level, Wikström and Wikström (2005) have found that private schools heavily inflate grades.

Evidence showing that the same is happening in compulsory schools is scarcer. However, the bias (i.e. difference between grades and national tests results) recorded in my data is larger at private schools compared to public schools. The mean bias for private schools and public schools is 0.48 points and 0.23 points, respectively. Looking only at mathematics, the mean bias for private and public schools are 1.31 points and 1.14 points, respectively. Consequently, special attention will be devoted to the effect of certified teachers in private and public schools.

Specification 3, which allows for heterogeneous effects of certified teachers across subjects, has been run on private and public schools separately. The outcomes of interest are the same as in the previous section and the results can be viewed in Table 5. Note that there are only 252 private schools in my sample, leading the estimates for private schools to be more imprecise in comparison to the estimates for public schools.

Comparing private and public schools, there are no significant differences in the effect of certified teachers on student achievement in English or Swedish. However, certified mathematics teachers have a larger, positive effect on average grades in private schools. In fact, the effect on public schools are not even significantly different from zero. Hence, the positive effect seen in the main results (Table 4) of mathematics teachers on grade, is primarily driven by private schools. A similar pattern emerges when looking at the share of students receiving a passing grade in mathematics[12]. Turning to standardized national tests, the gain for private schools in mathematics is significantly larger than for public schools. Lastly, the negative effect of certified mathematics teachers on bias is larger for private schools. However, the difference between the two estimates (-1.264 and -1.005) is not significantly different at any conventional level.

In summary, private schools see larger effects of certified mathematics teachers on student achievement. The explanation behind these results cannot be extracted directly from this analysis. However, it is possible that private schools are better at placing certified mathematics teachers in the classrooms where they have the highest impact, or that high achieving students are more affected by certified mathematics teachers (private schools score higher on national tests in mathematics compared to public schools). A rationale supporting the latter explanation would be that students with an initial high level in mathematics knowledge, are in need of teachers with higher levels of knowledge to develop their mathematical thinking positively. From an equity perspective, it is worrisome that private schools, which already teach more privileged students are disproportionately

[12]However, this difference between private and public schools is not statistically significant at the 5%-level.

Table 5: Effect of teacher certification on student achievement in core subjects (mathematics, English, and Swedish) in private and public schools, respectively

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Grade		Pass		National test		Bias	
Variables	Private	Public	Private	Public	Private	Public	Private	Public
Share of certified teachers	0.867***	0.163	0.0322***	0.0122*	2.205***	1.182***	-1.264***	-1.005***
	(0.210)	(0.125)	(0.00828)	(0.00642)	(0.364)	(0.164)	(0.285)	(0.150)
Interaction: Certified * English	-0.879***	-0.00131	-0.0345***	-0.00420	-2.359***	-1.179***	1.311***	1.120***
	(0.255)	(0.147)	(0.00896)	(0.00704)	(0.374)	(0.192)	(0.293)	(0.176)
Interaction: Certified * Swedish	-1.402***	-0.646***	-0.0394***	-0.0389***	-2.801***	-1.427***	1.414***	0.806***
	(0.248)	(0.166)	(0.00839)	(0.00851)	(0.386)	(0.211)	(0.310)	(0.182)
F-test (p-value): English	0.9373	0.1263	0.6083	0.0868	0.2875	0.9728	0.7366	0.2847
F-test (p-value): Swedish	0.0043	0.0003	0.1365	0.0001	0.0036	0.0818	0.3747	0.1020
Observations	3,701	11,492	3,662	11,482	3,287	10,410	3,274	10,301
R-squared	0.686	0.616	0.204	0.365	0.771	0.783	0.393	0.528
School-by-year FE	x	x	x	x	x	x	x	x
Subject FE	x	x	x	x	x	x	x	x

Notes: National test vary between 0-20 points with 10 points corresponding to a passing grade. Bias is constructed as $Grade - National\,test$. *Private* refers to schools operated by non-profit organizations or profit-driven companies, whereas *Public* refers to schools operated by a municipality. *F-test (p-value): English* is the p–value of the F-test testing the null hypothesis that the sum of the coefficients of the share of certified teachers and the interaction between certified and English equals zero $((\alpha_1 + \alpha_2) = 0)$. *F-test (p-value): Swedish* displays the equivalent value but for the subject Swedish $((\alpha_1 + \alpha_3) = 0)$. Data on national test results in mathematics are missing for 2017. Cluster-robust standard errors are in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

rewarded in a shift towards more certified teachers. From a fairness perspective, on the other hand, certified teachers can counteract the grade inflation which is especially pronounced for private schools.

7.2 Female and male student achievement

In much of the Western world, concerns have been raised regarding the fact that girls are outperforming boys in school. Also, the PISA surveys have revealed an increasing divide between the genders. Thus, it is of interest to understand how certified teachers affect the achievement of these two groups of students. Hence, I have run Specification 3 in two versions, using different outcome variables. One outcome variable is the average female national test results and one is the average male national test results. The results can be seen in Table 6, Columns 1 and 2 and I find no significant differences between the genders.

It is notable that the coefficient on the share of certified teachers (in mathematics) is lower for both females and males, compared to when the aggregated measure of national test results is used. The reason for this is that a restricted sample is used for this analysis. My data include average national test results for each school, divided into female and male averages. However, data is not made public if the observation is based on less than ten students. Consequently, schools that display data for both genders together since they have more than ten ninth graders in total, might not display gender segregated data because they have less than ten students of either gender. This leads to the exclusion of the smallest schools in this analysis. To see the main results for this restricted sample (where schools that do not have gender segregated national tests data are removed), see Table A4 in the Appendix.

7.3 Large and small schools

The size of Swedish compulsory schools varies significantly. In my sample, the smallest school has only 14 students, whereas the largest has 1,300. Larger schools are typically located in urban areas where also the students exhibit other characteristics. To test whether the size of the school interacts with the effect of certified teachers, the sample has been divided into schools larger than 300 students and schools smaller than or equal to 300 students (on average across the five years). The results of the two separate regressions can be seen in Table 6, Columns 3 and 4. The effect of teacher certification in mathematics at large schools is 1.514 points whereas it is 1.466 in small schools. I cannot reject the null hypothesis that the effects across large and small schools are the same. It is worth noting that small schools are more prone to measurement errors in the explanatory variable. Since the smallest schools only have a few teachers, measurement errors in the number of certified teachers result in large variation in the share of certified teachers. Hence, much of the variation recorded in the share of certified teachers for smaller schools, could be the result of measurement errors.

7.4 Students with non-Swedish backgrounds

Increased housing segregation between Swedish natives and immigrants, especially non-Europeans, has been a growing concern and is well-documented (see e.g. Hårsman and Quigley, 1995; Murdie and Borgegard, 1998). Since compulsory school allocation is primarily based on where the students live, segregation becomes a concern for the schools as well. 805 schools in the sample (corresponding to 78%) have more than 10% of their students coming from non-Swedish backgrounds, when taking the average across the five years (2014-2018). This is defined as the student not being born in Sweden, or that both parents are not born in Sweden. When separating these schools from the rest in the sample, I find no significant difference in the effect of certified mathematics teachers on national test results (Table 6, Columns 5 and 6).

Table 6: Effect of teacher certification on national test results (0-20 points) in core subjects (mathematics, English, and Swedish) for different sub-samples of schools

Variables	(1) Females	(2) Males	(3) Large	(4) Small	(5) Immigrants: $> 10\%$	(6) Immigrants: $\leq 10\%$	(7) Educated: $> 50\%$	(8) Educated: $\leq 50\%$
Share of certified teachers	1.270***	1.334***	1.514***	1.466***	1.412***	0.981***	1.272***	1.409***
	(0.171)	(0.171)	(0.217)	(0.217)	(0.179)	(0.328)	(0.225)	(0.227)
Interaction: Certified * English	-1.380***	-1.367***	-1.389***	-1.706***	-1.429***	-1.339***	-1.344***	-1.568***
	(0.196)	(0.191)	(0.246)	(0.239)	(0.199)	(0.367)	(0.248)	(0.262)
Interaction: Certified * Swedish	-1.465***	-1.507***	-1.666***	-1.931***	-1.755***	-1.048***	-1.542***	-1.610***
	(0.207)	(0.202)	(0.264)	(0.254)	(0.212)	(0.403)	(0.259)	(0.279)
F-test (p-value): English	0.2836	0.7437	0.3446	0.0389	0.8716	0.0231	0.4754	0.2893
F-test (p-value): Swedish	0.1242	0.1790	0.4129	0.0014	0.0083	0.7884	0.0493	0.2818
Observations	12,053	12,053	7,223	6,474	10,542	3,155	7,556	6,141
R-squared	0.788	0.762	0.801	0.756	0.775	0.808	0.805	0.760
School-by-year FE	x	x	x	x	x	x	x	x
Subject FE	x	x	x	x	x	x	x	x

Notes: National test vary between 0-20 points with 10 points corresponding to a passing grade. Data on national test results in mathematics are missing for 2017. The variable measuring average national test results for females and males have fewer observations than the variable for both sexes, for integrity reasons. Large schools are defined as those with more than 300 students in total. The share of immigrants is the defined as the share of students not born in Sweden or with (both) parents not born in Sweden (at school-level). The share of educated is defined as the the share of students with at least one parent with tertiary education (at school-level). *F-test (p-value): English* is the p–value of the F-test testing the null hypothesis that the sum of the coefficients of the share of certified teachers and the interaction between certified and English equals zero ($(\alpha_1 + \alpha_2) = 0$). *F-test (p-value): Swedish* displays the equivalent value but for the subject Swedish ($(\alpha_1 + \alpha_3) = 0$). Cluster-robust standard errors are in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

7.5 Students with highly educated parents

As pointed out in the literature review, one of the main predictors of student outcomes is socioeconomic background. In my sample, schools where more than half of the students have at least one parent with tertiary education score 12.1 points on the national tests in mathematics, on average, whereas the schools with fewer students with highly educated parents score 9.9 points. When running Specification 3 on these two groups of schools separately, the effect of certified mathematics teachers on national test does not differ significantly between them (Table 6).

8 Robustness of the results

This section tests the robustness of the main results by exploring potential confounders of the estimates.

8.1 Spillover effects of teachers across subjects

As previously discussed, the main specifications rely on the assumption that there are no spillover effects of certified teachers across subject. Theoretically, such spillover effects could be both positive and negative. To test whether there are any spillover effects, I will use school fixed and year fixed effects, similar to Specification 1. However, I will allow for the possibility of certified teachers in other subjects to have an impact on the outcome variable. The proposed specification is the following:

$$Y_{stj} = \alpha_0 + \alpha_1 Certified_{stj} + \sum_{i \neq j} \alpha_i Certified_{sti} + \delta_s + \delta_t + \beta \mathbf{X}_{stj} + \varepsilon_{stj} \quad (4)$$

where Y_{stj} is the outcome of interest at school s in time t in subject j. $Certified_{stj}$ is the share of certified teachers in that subject and $Certified_{sti}$ is the share of certified teachers in another subject. δ_s refers to school fixed effects, δ_t refers to time fixed effects, and $\mathbf{X}_{stj}$ is a vector of controls including time-varying student characteristics at the school-level. ε_{stj} is the error term.

Table 7: The effect of certified teachers on national test results (0-20 points) per subject

Variables	(1) Math	(2) English	(3) Swedish
Share of certified mathematics teachers	0.472***	0.117	-0.152
	(0.168)	(0.107)	(0.127)
Share of certified English teachers	0.0185	0.0893	-0.115
	(0.152)	(0.0840)	(0.0903)
Share of certified Swedish teachers	-0.113	-0.0418	-0.0299
	(0.174)	(0.0993)	(0.137)
Observations	3,436	4,214	4,013
R-squared	0.289	0.031	0.058
Year FE	x	x	x
School FE	x	x	x
Controls	x	x	x

Notes: National test vary between 0-20 points with 10 points corresponding to a passing grade. Data on national test results in mathematics are missing for 2017. The controls include the share of female students in ninth grade, the share of students not born in Sweden or with (both) parents not born in Sweden (at school-level), the share of students with at least one parent with tertiary education (at school-level), and the number of ninth graders taking the subject in question. Cluster-robust standard errors are in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

When I run this specification, the three core subjects are included, and the outcome variables are the national test result in each subject. If there are spillover effects across subjects, any α_i should be different from zero. Looking at Table 7, Column 1, the variables of interest are the share of certified teachers in English and Swedish. None of these coefficients is statistically significantly different from zero, indicating that certified teachers in English and Swedish have no effect on national test results in mathematics. The same conclusion is valid for English and Swedish, where certified teachers in other subjects have no significant effect on national test results. Although this specification could be subject to biases due to failures to control for time-varying unobservables, these results indicate that spillover effects are not biasing the main results of this paper.

8.2 The effect of certified teachers on number of students

As pointed out in Section 4.1, not all students take Swedish as a subject. Rather, students who are not native can take "Swedish as a second language". It is up to teacher's discretion to decide whether a student should take regular Swedish or "Swedish as a second language". If certified and non-certified teachers make different decisions, on average, it could bias the estimated coefficients, since this study uses aggregated data for each school. An example of this would be that non-certified teachers think that more students should take "Swedish as a second language", in practice removing low-performing students from my data. This in turn would lead schools with more non-certified teachers to have higher average national test results in Swedish.

Hence, I want to find out if there is any support for this hypothesis by examining the number of ninth graders recorded within each of the core subjects. If fewer students take Swedish when the share of certified Swedish teachers increase, while no such pattern can be seen within mathematics or English, this would indicate that certified and non–certified Swedish teachers make different judgements regarding whether a student should take Swedish or "Swedish as second language". Thus, Specification 1 has been run with the outcome variable being the number of students in ninth grade that take the subject in question and the vector of controls removed. Moreover, the primary group of students that take "Swedish as second language" are those with a non-Swedish background. Hence, I have split the sample depending on the share of students with non-Swedish backgrounds, with 10% as the limit.

Table 8: Effect of teacher certification on number of ninth graders taking each subject

	(1)	(2)	(3)	(4)	(5)	(6)
	Immigrants: $> 10\%$			Immigrants: $\leq 10\%$		
VARIABLES	Math	English	Swedish	Math	English	Swedish
Share of certified teachers	-1.710	0.289	3.781***	1.491	-0.437	2.602
	(1.398)	(1.007)	(1.245)	(3.230)	(1.647)	(2.862)
Observations	3,950	3,932	3,857	1,147	1,142	1,169
R-squared	0.098	0.072	0.052	0.128	0.110	0.112
Year FE	x	x	x	x	x	x
School FE	x	x	x	x	x	x

Notes: Cluster-robust standard errors are in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Looking at the results in Table 8, only within Swedish and at schools with at least 10% non-native students, the share of certified teachers has a significant effect on the number of students recorded in the grade data for each subject. The estimate suggests that going from none to all certified Swedish teachers leads to that almost four more students in the ninth grade are included in the student achievement data for Swedish. Although this finding supports the hypothesis that certified Swedish teachers think that fewer students need to take "Swedish as a second language", I have no conclusive evidence that so is the case. Regardless, I caution against interpreting the negative coefficient of certified Swedish teachers on student achievement as causal, since it is possible that the negative effect is driven by an added number of low-performing students, rather than a deterioration in student achievement at the individual level.

8.3 The exclusion of confounding observations

This sub-section examines potential reasons for confounded estimates by excluding certain observations from the original sample while sticking to my preferred specification (3). The effects of certified teachers on national test results and bias can be seen in Table 9, with parts of the original sample removed. None of these exclusions change the main conclusions qualitatively: The effect of certified mathematics teachers on national test results is constantly positive and substantial, whereas the effect of English and Swedish is slightly negative, sometimes significantly so. But since the magnitude of the coefficients are sensitive to these exclusions, I caution against drawing far-reaching conclusions based on the estimated magnitudes of this thesis. The rest of this sub-section will be devoted to explaining why certain observations have been excluded in Table 9.

In Columns 1 and 2, the subject Swedish has been excluded. This is due to the problems with using aggregated student achievement in Swedish that was discussed in Section 8.2.

In Columns 3 and 4, all observations from academic year 2014/2015 has been excluded. One major concern throughout this study is that obtaining a certification is not equivalent to obtaining a degree or taking additional courses, since the teacher has to actively apply for the certification rather than it being automatically granted upon graduation or finished additional coursework. That is, when a school increases its share of certified teachers between two years, it is not certain that it has employed new teachers that are certified or that its existing teachers have taken courses or finished their degrees. It could simply be the case that some of its teachers applied for a certification that they already were qualified for. In that case, one more certified teacher is just representing a piece of paper at a cost of SEK 1,500. The risk of the data primarily picking up applications for certification from already qualified teachers, is the greatest in academic year 2014/2015. In this year, the teacher certification reform was not fully implemented, and many non-certified teachers were able to keep on teaching without any restrictions. Hence, I have excluded all observations from that year.

In Columns 5 and 6, all schools with students in grade five or lower have been excluded. A problem with using data at the school-level is that it does not contain information on which students the teachers are assigned to. Even if a school increases its share of certified teachers, those certified teachers could be assigned to students in grades other than the ninth grade. In that case, the coefficient on certified teachers should equal zero. Although I cannot track which classes the certified teachers are assigned to, I can increase the probability of them being assigned to the ninth grade by excluding schools that teach the largest number of grades. Compulsory school in Sweden includes ten years (pre-school to ninth grade). However, it is common that schools divide the grades between them in such a way that students need to change school once or twice during their time in compulsory school. It is common to switch schools between third and fourth grade, and between sixth and seventh grade. For those switching only once, it is also common to switch between fifth and sixth grade. This means that some of the schools in my sample teach all grades through compulsory schools, whereas others teach only from grade six or seven, and up. Consequently, I have only included the schools that teach students in grade six and higher.

In Columns 7 and 8, schools with 75% certified teachers or less on average in the core subjects have been excluded. When interpreting the estimates, I have assumed that the effect of certified teachers is linear and independent of the previous share of certified teachers. As an example, increasing the share of certified teachers by 10 percentage points increase national test results by

Table 9: Effect of teacher certification on student achievement in core subjects (mathematics, English, and Swedish) with certain observations excluded

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Swedish excluded		2014 excluded		$<$ Grade 6 excluded		$\leq$ 75% certified excluded	
Variables	National test	Bias	National test	Bias	National test	Bias	National test	Bias
Share of certified teachers	1.303***	-1.199***	1.422***	-1.022***	1.093***	-0.814***	0.929***	-0.675***
	(0.163)	(0.153)	(0.162)	(0.143)	(0.195)	(0.175)	(0.294)	(0.251)
Interaction: Certified * English	-1.638***	1.308***	-1.437***	0.985***	-1.174***	0.788***	-1.172***	0.523*
	(0.196)	(0.173)	(0.182)	(0.156)	(0.220)	(0.196)	(0.304)	(0.283)
Interaction: Certified * Swedish			-1.948***	0.786***	-1.023***	0.558***	-0.891**	0.521*
			(0.200)	(0.175)	(0.234)	(0.210)	(0.360)	(0.312)
F-test (p-value): English	0.0092	0.3400	0.8757	0.6758	0.4915	0.8301	0.0537	0.3043
F-test (p-value): Swedish			0.0002	0.0597	0.6248	0.0652	0.8574	0.4773
Observations	8,998	8,880	10,733	10,729	6,270	6,200	6,281	6,220
R-squared	0.868	0.603	0.777	0.508	0.799	0.550	0.794	0.499
School-by-year FE	x	x	x	x	x	x	x	x
Subject FE	x	x	x	x	x	x	x	x

Notes: National test vary between 0-20 points with 10 points corresponding to a passing grade. Bias is constructed as $Grade - National\,test$. Data on national test results in mathematics are missing for 2017. In Columns 1-2, all observations concerning the subject Swedish are excluded. In Columns 3-4, all observations from academic year 2014/2015 are excluded. In Columns 5-6, all schools that have any students in grades 1-5 are excluded. In Columns 7-8, all schools that have less than 75 percent certified teachers in the core subjects, on average, are excluded. *F-test (p-value): English* is the p–value of the F-test testing the null hypothesis that the sum of the coefficients of the share of certified teachers and the interaction between certified and English equals zero $((\alpha_1 + \alpha_2) = 0)$. *F-test (p-value): Swedish* displays the equivalent value but for the subject Swedish $((\alpha_1 + \alpha_3) = 0)$. Cluster-robust standard errors are in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

0.0254 points and increasing certification by 100 percentage points increase results by 0.254 points. Moreover, I assume that increasing certified teachers by 10 percentage points always leads to increased results of 0.0254 points, regardless if the school went from 0 to 10 percent certification or from 90 percent to 100 percent certification. This interpretation has been made despite the fact that the estimates are based mostly on small changes in certification share and that schools generally started out with the majority of teachers in core subjects being certified. Moreover, the administrative burden put on certified teachers depends on the ratio of non–certified to certified teachers. The mean year-on-year change in certified share is 15 percentage points and the mean certified share of teachers (across all years) is 72.5 percent. To test the assumptions of linearity and that the initial level of certification is without effect, I have run Specification 3 with only the schools that have more than 75 percent certified teachers, on average across all years. These schools not only started out with more certified teachers, but have had smaller year-on-year changes in certification share (12 percentage points on average). It should be noted that this test is not optimal for my purposes, since these schools can be different from the rest of the sample in other regards. It is possible that those differing factors interact with the share of certified teachers, creating other results.

8.4 The use of weighted least squares

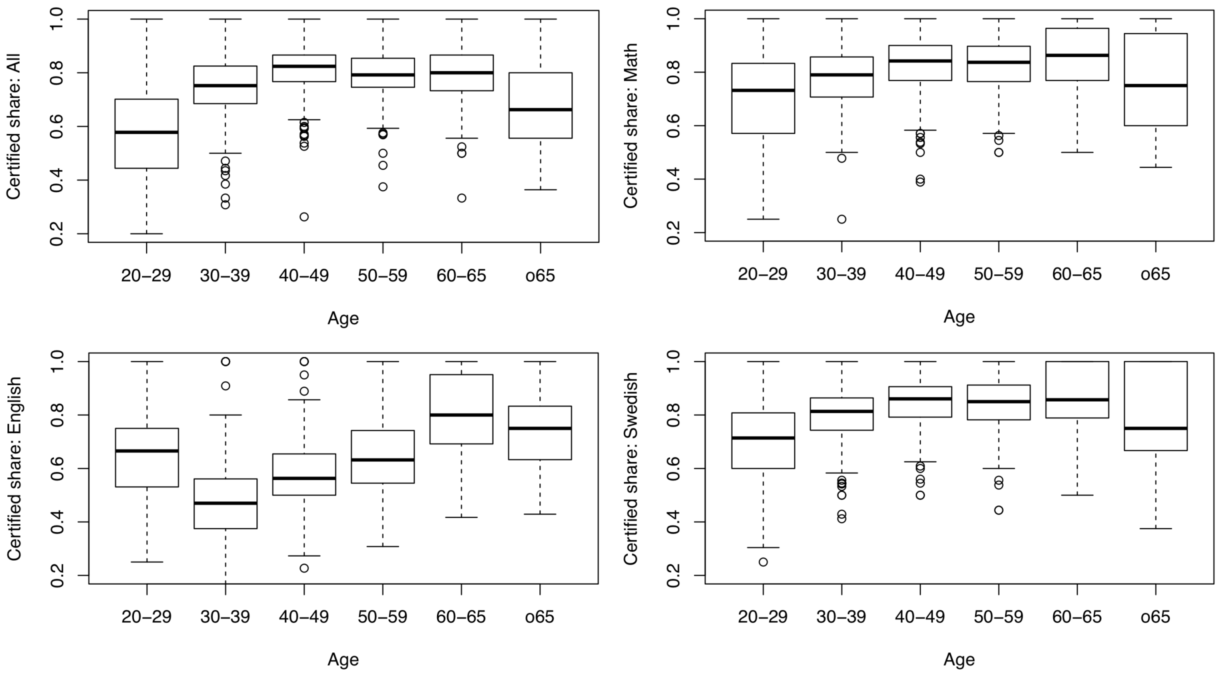
Since the outcome variables are averages of differing numbers of ninth graders for each school and each year, it could be argued that weighted least squares should have been the preferred estimation method, rather than ordinary least squares. One such argument is that weighting could correct for student-size related heteroscedasticity in the error terms. Given, that I am already applying heteroscedasticity-robust standard errors, I have opted to use ordinary least squares as my preferred estimation method. However, I have included the results of using weighted least squares, which are similar to the main results (Table 4), in the Appendix, Table A5.

8.5 Correlation between teacher certification and experience

As mentioned in the literature review, authors such as Krueger (1999) and Rivkin, Hanushek, and Kain (2005) have found positive effects of teacher experience on test results, at least in the initial years of teaching. One might worry that teacher experience and teacher certification is positively correlated, pushing the estimated effects of teacher certification on national test results upwards. If data on average teacher experience for each subject was available, it could simply be added as a control to avoid this type of omitted variable bias. However, data on neither experience nor age is available at the school-level. But the share of certified teachers in each age group in 2018 has been made available at the municipality-level.

Figure 3 displays four boxplots, one of all teachers and three of teachers in each core subject (mathematics, English, and Swedish). Looking at the first boxplot displaying all teachers, those aged 20-29 are certified to a lesser extent. This is reasonable given that many substitute teachers are young and non-certified. Additionally, some work as teachers before deciding whether to pursue a pedagogical degree. Also at the upper end of the age distribution, those aged over 65 are less likely to be certified. It is easy to imagine that elderly teachers without a certification decided against pursuing a lengthy degree when the certification requirement was announced. Most importantly, the median share of certified teachers in the ages 30-65 is evenly spread across the age groups. Looking at the teacher population in Sweden, the vast majority of teachers are aged 30 to 65. The pattern just described, is similar when looking specifically at teachers in Mathematics and Swedish. Hence, correlation between experience and certification should not present an issue. However, a different pattern can be seen for English teachers, where age correlates positively with certification. On the other hand, I have not detected any significantly positive effects of certified teachers in English.

Figure 3: Boxplots of the share of certified teachers per subject and age group in academic year 2018/2019



Notes: *All* refers to teachers across the entire school, i.e. all 16-18 subjects in compulsory school. The plots are created with data on the municipality-level. Source: Author's own rendering of data from the Swedish National Agency for Education (2020b).

9 Discussion

To briefly summarize the results, I find that having certified teachers in mathematics significantly increase student achievement in terms of average grade, the share of students passing the subject, and average national test results. The positive effects on national test results are robust in a series of robustness tests. The effects are greater on national tests, compared to grades, leading to positive outcomes looking from a fairness perspective. In English, I find no significant effects of certified teachers whereas in Swedish, certified teachers have negative effects on average student outcomes. From an analysis of the correlation between the number of students taking each subject and the share of certified teachers, I have concluded that these negative results likely stem from the usage of aggregated, rather than individual-level, data.

9.1 Internal validity

If the empirical strategy is to provide causal estimates of the effect of teacher certification, the identifying assumptions need to hold. Specifications 2 and 3 assume that there are no spillover effects of certified teachers across subjects, and that there are no time-varying variables at the subject-level that correlate with both the share of certified teachers and the outcome variable in that subject. The assumption of no spillover effects was tested in Section 8.1, not finding any evidence in favor of any spillovers. However, the test used a version of Specification 1 (i.e. school fixed effects, time fixed effects, and a vector of controls), which could also be subject to bias. As previously stated, this is reasonable given that the vector of controls is not rich, including only four variables. Hence, there are only indicative evidence of that the identifying assumption of no spillovers holds. The assumption of no subject-specific and time-varying variables that lead to omitted variables bias could also be questioned. An example of a variable that possibly leads to omitted variables bias is the level of teacher experience within each subject. Although I have been unable to include this variable in my main specification, the correlation between teacher certification and age has been explored in Section 8.5. I deem that the failure to control for teacher experience does not pose a serious threat to the identifying assumptions, but it is possible that there are other variables that do.

The main threat to the validity of my research design, is a lack of data. Most importantly, without student-teacher matched data, my analysis lacks the sufficient level of detail. With the current data, some of the certified teachers that are added to schools never teach any ninth graders. Consequently, the magnitudes of my estimates only say something about what happens with the student achievement of ninth graders when the share of certified teachers is increased at the school-level. With student-teacher matched data, it would be possible to estimate what happens to a single class of ninth graders when they go from having a non-certified to a certified teacher. Another problem is that I cannot identify if an increase in certified teachers is due to an actual change in the educational level of the teachers or if teachers that had already met the requirements for a certification applied at a later point in time. Also, no distinction has been made between non–certified teachers and teachers certified in other subjects than the ones they are teaching in. Moreover, when data is aggregated, the means can be manipulated not only by making individual students perform better, but by adding or removing high/low performers from the school. In the case of the subject Swedish, this can be done without the students having to change school, leading the estimates to capture both the effect on achievement and the increase in low performing students that seems to come with having more certified teachers. Further examples of lacking data is the missing values of national test results in academic year 2017/2018 and the availability of control variables in Specification 1.

Simultaneity errors, or reverse causation, should not pose an issue. The share of certified teachers is recorded in the beginning of the academic year (in October), whereas the national tests are taken and grades are granted towards the end of the academic year. Thus, teachers cannot observe "current" (or rather future) student achievement before they decide where to teach.

In summary, the lack of more fine-grained data is the greatest threat to internal validity. The direction of the effect of certified teachers is robust across the thesis, but the magnitudes of the effects are not. Consequently, I caution against interpreting the magnitude of the estimates as the true coefficients, or drawing far-reaching conclusions based on them.

9.2 External validity

The results of this thesis are unlikely to hold up in other settings, in terms of both time and space. To begin with, the Swedish institutional context sets it apart from other places in the world, even from other industrialized nations. The extensive welfare provided in Sweden has especially large implications for the Swedish schools. As an example, few countries in the world fund private schools with such generous amounts that they can be highly profitable, without having admission fees. Another distinctive feature of the Swedish context is the high share of immigrants in the general population combined with increased housing segregation, which leads to schooling segregation. Additionally, the interest for the teaching profession has dwindled in the past decades, leading many schools to struggle to hire enough teachers. All of these factors likely contribute to how certified teachers are hired, what kind of schools and with what means the schools compete for certified teachers, and ultimately how effective certified teachers are. Additionally, what constitutes a certified teacher is specific to Sweden.

Furthermore, these results may not be valid for a longer time period in Sweden. I want to stress that this thesis has studied how effective certified teachers are, in comparison to non-certified teachers. Ultimately, the results will depend on both the population of certified teachers and the population of non-certified teachers. Little attention has been paid to the latter group, mainly because knowledge and data about the non-certified teachers are scarce. Depending on a vast range of factors such as unemployment rates and policy changes, the population of non-certified teacher could change dramatically going forward. Even if the definition of teacher certification remains the same and the same type of individuals become certified teachers, this would have an impact on the estimated effect of certified teachers. Consequently, my estimates of the effects of certified teacher are specific to this moment in time.

10 Conclusion

In the beginning of this thesis, I set out to study the effects of teacher certification on ninth grade student achievement in Sweden, where a national, subject-specific teacher certification was introduced in 2011. By applying panel data spanning over five academic years, 2014/2015-2018/2019, to fixed effects models, I conclude that certified teachers within mathematics have significant and positive effects on student achievement, at least in terms of standardized national test results. However, I do not find any positive effects of certified teachers in the two other core subjects studied, namely English and Swedish. These results are in line with the previous literature, taking place primarily in the US, studying teacher certification and degrees.

Using my preferred specification with school-by-year fixed effects, I find that certified teachers in mathematics increase the average grade of ninth graders, the share of ninth graders passing the subject, and the average standardized national test result. Moreover, the increase in national test results is substantially larger than the increase in grades, meaning that these teachers put a downward pressure on grade inflation. This is encouraging, given that among the three core subjects, mathematics is the one with the largest discrepancy between grades and standardized national test results. That is, mathematics teachers in Sweden consistently grant grades higher than the true knowledge level of their students.

In English, I find no significant effects of having certified teachers. In Swedish, I find negative effects of certified teachers when applying the preferred specification. This negative effect is persistent throughout a selection of the robustness tests. Moreover, I find indicative evidence of that it is a result of using school-level, rather than individual-level data. In compulsory school, Swedish is the only subject where the students are divided into different tracks based on the level of the students. A minority of the students take the lower level track called "Swedish as second language" and the student achievement within this track is not recorded in my study. It is up to teacher's discretion to decide which students should take which track. I find that the share of certified teachers within schools, correlate positively and significantly with the number of students taking the regular track ("Swedish"). This implies that certified teachers deem that fewer students require the lower track "Swedish as second language". This non-random inclusion of students taking the regular track in schools with more certified Swedish teachers, likely decreases the mean student achievement.

When studying the heterogeneous effects of certified teachers across different types of schools and students, I find no heterogeneity with respect to gender, school size, the share of immigrant students at the schools, and the share of students with highly educated parents. On the other hand, I find that private schools see a significantly stronger effect of certified mathematics teachers on national test results compared to public schools.

The result that certified teachers in mathematics have positive effects on standardized national tests results, is robust to the exclusion of possibly confounding observations and a model with school and year fixed effects (Specification 1). However, the positive effect of certified mathematics teachers on grades and the share of students passing mathematics is not robust to that model. Furthermore, this thesis faces several limitations, why I am conservative in the interpretation of the magnitudes of the effects of certified mathematics teachers. The main limitations to internal validity includes the lack of individual-level data, the lack of data that are matched to cohorts within schools, and the lack of information about when increases in the share of certified teachers are due to that teachers have just finished the required courses and when it is due to that teachers have simply applied for the certification (although they were eligible all along). Most importantly, no further understanding of the mechanisms behind these results has been developed through my study.

This thesis is the first to study the effect of teacher certification on student achievement after the introduction of the national teacher certification in Sweden and the results have important policy implications. Based on this study, having certified teachers, i.e. teachers with a pedagogical degree and subject-specific knowledge, significantly increases student achievement in mathematics. However, no such effects can be discerned in the subjects English or Swedish. Speaking purely from the perspective of student achievement, the requirement for schools to have certified teachers in all subjects, risks discouraging non-certified but highly qualified teachers from entering or remaining within the teaching profession. It is possible that other ways to encourage pedagogical degrees for mathematics teachers specifically, would have been more effective in improving student achievement at the compulsory schooling level. It should also be noted that the mean share of certified mathematics teachers in Swedish schools has not increased during the five years studied. When it comes to narrowing the gap between top- and bottom-performing schools, I find no evidence in favor of certified teachers. That is not to say that the teacher certification has not had other beneficial outcomes that are beyond the scope of this thesis.

Naturally, policy makers are in need of more robust results than presented in this study and further research on teacher certification within the Swedish context is encouraged. First, only three subjects have been the focus of this study. It is possible that other subjects such as science are positively affected by teacher certification. Second, I implore future researchers to study the effects of teacher certification on the individual student-level. By doing so, studies can be done on the effect of certified teachers on certain types of students, rather than certain types of schools. Lastly, the mechanisms behind the effects of teacher certification in Sweden have not been within the scope of this thesis but are crucial to understand in order to design effective policy measures.

References

Andersson, C., & Waldenström, N. (2007). *Teacher certification and student achievement in Swedish compulsory schools.* Institute for evaluation of labour market and education policy.

Angrist, J. D., & Lavy, V. (1999). Using maimonides' rule to estimate the effect of class size on scholastic achievement. *The quarterly journal of economics*, *114*(2), 533–575.

Anthony, E., & Goldhaber, D. D. (2007). Can teacher quality be effectively assessed? National board certification as a signal of effective teaching. *The review of economics and statistics*, *89*(1), 134–150.

Båvner, P., Barklund, A., Hellewell, A., & Svensson, M. (2011). OECD-overcoming school failure. Country background report Sweden. *Ministry of education and research*, *46*.

Cantrell, S., Fullerton, J., Kane, T. J., & Staiger, D. O. (2008). *National board certification and teacher effectiveness: Evidence from a random assignment experiment.* National bureau of economic research.

Chingos, M. M., & Peterson, P. E. (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), 449–465.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). *How and why do teacher credentials matter for student achievement?* National bureau of economic research.

Coenen, J., Cornelisz, I., Groot, W., Maassen van den Brink, H., & Van Klaveren, C. (2018). Teacher characteristics and their effects on student test scores: A systematic review. *Journal of economic surveys*, *32*(3), 848–877.

Cowan, J., & Goldhaber, D. (2016). National board certification and teacher effectiveness: Evidence from Washington state. *Journal of research on educational effectiveness*, *9*(3), 233–258.

Dewey, J., Husted, T. A., & Kenny, L. W. (2000). The ineffectiveness of school inputs: A product of misspecification? *Economics of education review*, *19*(1), 27–45.

Glazerman, S., Mayer, D., & Decker, P. (2006). Alternative routes to teaching: The impacts of teach for america on student achievement and other outcomes. *Journal of policy analysis and management: The journal of the association for public policy analysis and management*, *25*(1), 75–96.

Goldhaber, D. D., & Brewer, D. J. (1997a). *Evaluating the effect of teacher degree level on educational performance.* Developments in school finance, 1996.

Goldhaber, D. D., & Brewer, D. J. (1997b). Why don't schools and teachers seem to matter? Assessing the impact of unobservables on educational productivity. *Journal of human resources*, 505–523.

Goldhaber, D. D., & Brewer, D. J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational evaluation and policy analysis*, *22*(2), 129–145.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of economic literature*, *24*(3), 1141–1177.

Harris, D. N., & Sass, T. R. (2009). The effects of NBPTS-certified teachers on student achievement. *Journal of policy analysis and management: The journal of the association for public policy analysis and management*, *28*(1), 55–80.

Hårsman, B., & Quigley, J. M. (1995). The spatial segregation of ethnic and demographic groups: Comparative evidence from stockholm and san francisco. *Journal of urban economics*, *37*(1), 1–16.

Hawk, P. P., Coble, C. R., & Swanson, M. (1985). Certification: It does matter. *Journal of teacher education*, *36*(3), 13–15.

Horoi, I., & Bhai, M. (2018). New evidence on national board certification as a signal of teacher quality. *Economic inquiry*, *56*(2), 1185–1201.

Ingersoll, R., & Gruber, K. (1996). *Out-of-field teaching and educational equality.* Washington, DC: U.S. department of education.

Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of education review*, *27*(6), 615–631.

Krueger, A. B. (1999). Experimental estimates of education production functions. *The quarterly journal of economics*, *114*(2), 497–532.

Monk, D. H., & King, J. A. (1994). Multilevel teacher resource effects in pupil performance in secondary mathematics and science: The case of teacher subject matter preparation. *Choices and consequences: Contemporary policy issues in education*, 29–58.

Murdie, R. A., & Borgegard, L.-E. (1998). Immigration, spatial segregation and housing segmentation of immigrants in metropolitan Stockholm, 1960-95. *Urban studies*, *35*(10), 1869–1888.

Myrberg, E. (2007). The effect of formal teacher education on reading achievement of 3rd-grade students in public and independent schools in Sweden. *Educational Studies*, *33*(2), 145–162.

Organisation for Economic Co–operation and Development. (2010). *Pisa 2009 results: Overcoming social background: Equity in learning opportunities and outcomes (volume ii).* OECD publishing Paris.

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

Sahlgren, G. H. (2016). *Regulation and funding of independent schools: Lessons from Sweden.* Fraser Institute.

Statistics Sweden. (2019). Behöriga förstahandssökande och antagna till program efter inriktning höstterminerna 1998-2019.

Swedish Government. (2008). Proposition 2008/09:66: En ny betygsskala.

Swedish Government. (2010). Proposition 2010/11:20: Legitimation för lärare och förskollärare.

Swedish National Agency for Education. (2020a). Nationella prov i grundskolan [Accessed: 2020-04-15]. https://www.skolverket.se/undervisning/grundskolan/nationella-prov-i-grundskolan

Swedish National Agency for Education. (2020b). Sök statistik [Accessed: 2020-02-15]. https://www.skolverket.se/skolutveckling/statistik/sok-statistik-om-forskola-skola-och-vuxenutbildning?sok=SokD

Taube, K., Fredriksson, U., Rasmusson, M., Karlsson, K., Pettersson, A., Ingemansson, I., Hammarberg, N., Oscarsson, M., & Sundgren, M. (2010). *Rustad att möta framtiden?: Pisa 2009 om 15-åringars läsförståelse och kunskaper i matematik och naturvetenskap*. Swedish national agency for education.

Wayne, A. J. (2002). Teacher inequality. *Education policy analysis archives*, *10*, 30.

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of educational research*, *73*(1), 89–122.

Wikström, C., & Wikström, M. (2005). Grade inflation and school competition: An empirical analysis based on the swedish upper secondary schools. *Economics of education review*, *24*(3), 309–322.

Appendix

Additional results of Specification 1

Table A1: Effect of certified teachers on grade (0-20 points) per subject

Variables	(1) Math	(2) Math	(3) English	(4) English	(5) Swedish	(6) Swedish
Share of certified teachers	0.00313	-0.0609	-0.0561	0.0165	-0.147	-0.125
	(0.116)	(0.115)	(0.0724)	(0.0752)	(0.0919)	(0.0970)
Observations	5,097	4,436	5,074	4,447	5,022	4,192
R-squared	0.076	0.121	0.034	0.067	0.014	0.049
Year FE	x	x	x	x	x	x
School FE	x	x	x	x	x	x
Controls		x		x		x

Notes: Grades vary between 0-20 points with 10 points corresponding to a passing grade. The controls include the share of female students in ninth grade, the share of students not born in Sweden or with (both) parents not born in Sweden (at school-level), the share of students with at least one parent with tertiary education (at school-level), and the number of ninth graders taking the subject in question. Cluster-robust standard errors are in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A2: Effects of certified teachers on the share of students receiving a passing grade per subject

Variables	(1) Math	(2) Math	(3) English	(4) English	(5) Swedish	(6) Swedish
Share of certified teachers	-0.0019	-0.0033	0.0030	0.0041	0.0016	0.0027
	(0.0070)	(0.0072)	(0.0035)	(0.0039)	(0.0035)	(0.0036)
Observations	5,077	4,436	5,061	4,447	5,006	4,192
R-squared	0.073	0.105	0.057	0.085	0.007	0.013
Year FE	x	x	x	x	x	x
School FE	x	x	x	x	x	x
Controls		x		x		x

Notes: Grades vary between 0-20 points with 10 points corresponding to a passing grade. The controls include the share of female students in ninth grade, the share of students not born in Sweden or with (both) parents not born in Sweden (at school-level), the share of students with at least one parent with tertiary education (at school-level), and the number of ninth graders taking the subject in question. Cluster-robust standard errors are in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Effect of teacher certification on the difference between grade and national test result (bias) per subject

Variables	(1) Math	(2) Math	(3) English	(4) English	(5) Swedish	(6) Swedish
Share of certified teachers	-0.522***	-0.416***	-0.181**	-0.127	-0.0324	-0.0827
	(0.149)	(0.143)	(0.0760)	(0.0848)	(0.0937)	(0.111)
Observations	4,012	3,481	4,868	4,277	4,695	3,933
R-squared	0.159	0.177	0.015	0.023	0.026	0.034
Year FE	x	x	x	x	x	x
School FE	x	x	x	x	x	x
Controls		x		x		x

Notes: Bias is constructed as $Grade - National\ test$. Data on national test results in mathematics are missing for 2017. The controls include the share of female students in ninth grade, the share of students not born in Sweden or with (both) parents not born in Sweden (at school-level), the share of students with at least one parent with tertiary education (at school-level), and the number of ninth graders taking the subject in question. Cluster-robust standard errors are in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Main specifications (2 and 3) with restricted sample

Table A4: Effect of teacher certification on student achievement in the core subjects (mathematics, English, and Swedish)

Variables	(1) Grade	(2) Grade	(3) Pass	(4) Pass	(5) National test	(6) National test	(7) Bias	(8) Bias
Share of certified teachers	0.0167	0.354***	0.000507	0.0186***	0.266***	1.308***	-0.252***	-0.957***
	(0.0706)	(0.123)	(0.00313)	(0.00569)	(0.0777)	(0.159)	(0.0759)	(0.140)
Interaction: Certified * English		-0.352**		-0.0184***		-1.381***		1.028***
		(0.148)		(0.00647)		(0.180)		(0.156)
Interaction: Certified * Swedish		-0.693***		-0.0381***		-1.514***		0.842***
		(0.152)		(0.00741)		(0.191)		(0.165)
F-test (p-value): English		0.9866		0.9692		0.4330		0.4318
F-test (p-value): Swedish		0.0029		0.0005		0.0774		0.2798
Observations	11,943	11,943	11,943	11,943	12,053	12,053	11,943	11,943
R-squared	0.616	0.618	0.317	0.321	0.796	0.800	0.519	0.524
Year-School FE	x	x	x	x	x	x	x	x
Subject FE	x	x	x	x	x	x	x	x

Notes: Schools that do not present the national test results for each gender separately are excluded from this sample. Grade and National test vary between 0-20 points with 10 points corresponding to a passing grade. Bias is constructed as the difference between Grade and National test ($Grade - National\,test$). Data on national test results in mathematics are missing for 2017. *F-test (p-value): English* is the p–value of the F-test testing the null hypothesis that the sum of the coefficients of the share of certified teachers and the interaction between certified and English equals zero ($(\alpha_1 + \alpha_2) = 0$). *F-test (p-value): Swedish* displays the equivalent value but for the subject Swedish ($(\alpha_1 + \alpha_3) = 0$). Cluster-robust standard errors are in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Weighted regressions

Table A5: Effect of teacher certification on student achievement in the core subjects (mathematics, English, and Swedish)

Variables	(1) Grade	(2) Grade	(3) Pass	(4) Pass	(5) National test	(6) National test	(7) Bias	(8) Bias
Share of certified teachers	0.0781	0.380***	0.00238	0.0168***	0.323***	1.342***	-0.272***	-0.976***
	(0.0705)	(0.110)	(0.00310)	(0.00510)	(0.0792)	(0.157)	(0.0809)	(0.136)
Interaction: Certified * English		-0.277**		-0.0127**		-1.351***		1.039***
		(0.133)		(0.00558)		(0.171)		(0.153)
Interaction: Certified * Swedish		-0.734***		-0.0360***		-1.517***		0.861***
		(0.140)		(0.00715)		(0.186)		(0.157)
F-test (p-value): English		0.2641		0.2557		0.9128		0.5116
F-test (p-value): Swedish		0.0022		0.0008		0.1432		0.2969
Observations	15,193	15,193	15,144	15,144	13,697	13,697	13,575	13,575
R-squared	0.648	0.649	0.354	0.357	0.799	0.802	0.526	0.532
School-by-year FE	x	x	x	x	x	x	x	x
Subject FE	x	x	x	x	x	x	x	x

Notes: The coefficients are estimated with weighted least squares, with the number of ninth graders in each school as weights. Grade and National test vary between 0-20 points with 10 points corresponding to a passing grade. Bias is constructed as the difference between Grade and National test ($Grade - Nationaltest$). Data on national test results in mathematics are missing for 2017. *F-test (p-value): English* is the p–value of the F-test testing the null hypothesis that the sum of the coefficients of the share of certified teachers and the interaction between certified and English equals zero ($(\alpha_1 + \alpha_2) = 0$). *F-test (p-value): Swedish* displays the equivalent value but for the subject Swedish ($(\alpha_1 + \alpha_3) = 0$). Cluster-robust standard errors are in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.