BI Norwegian Business School - campus Oslo

GRA 19502

Master Thesis

Component of continuous assessment: Thesis Master of Science

Teacher Quality and High School Completion:

Evidence from Norwegian Register data

Navn:	Rosanna Johed, Synne Sandnes

Start:	02.03.2017 09.00
Finish:	01.09.2017 12.00

Rosanna N. I. Johed
Synne Sandnes

Teacher Quality and High School Completion: Evidence from Norwegian Register data

Hand-in date:
30.06.2017

Examination code and name:
GRA 19502 Master Thesis

Campus:
BI Oslo

Supervisor from BI: Jon H. Fiva

Co-supervisor from SSB: Trude Gunnes

Programme:
Master of Science in Business with Major in Economics

Content

Acknowledgements

This master thesis has been carried out during the autumn of 2016 and the spring 2017, as the final thesis for the master programme MSc in Business with a major in Economics at BI Norwegian Business School, in corporation with Statistics Norway (SSB).

The process has been interesting and educational, both academically and collaboratively.

We would like to express our gratitude to our supervisors Jon H. Fiva from BI and co-supervisor Trude Gunnes from Statistics Norway for useful comments and suggestions. We would also like to thank Sturla Løkken for help with STATA and support during the semester. The process would have been a lot harder and longer without your help. We are very grateful.

Additionally, we would like to thank the rest of the faculty at the Economics department at BI Norwegian Business School for guidance regarding both practical and thesis specific challenges.

Lastly, we thank SSB and Marte Rønning for welcoming us and giving us a great opportunity to get first-hand experience with how researchers work.

We hope you as a reader will find the topic and the thesis interesting.

Rosanna Johed and Synne Sandnes

Summary

Students assigned to high value-added teachers benefit in the long term. They are more likely to attend college, get higher income, and less likely to have children as teenagers (Chetty, Friedman, & Rockoff, 2011). Students who drop out of high school will be restricted from attaining higher education and thus face limited career possibilities later in life. Hence, the problem of high school dropouts is a major economic and social concern. In this thesis, we investigate whether middle school teachers education is a determining factor in the probability that students complete high school. We combine four data sets to estimate our main model with middle school fixed effects. The detailed data on teacher characteristics has been of particular interest, considering it has not been used before. The model has advantages as it uses the within group variation, specifically, within middle school variation in teacher education over time.

We focus on two key variables at middle school, teachers with and without teacher education, and teachers with short and long higher education. Specifically, we find that there is a positive relationship between middle school teachers with teacher education and high school completion. We find that a 10 percentage point increase in the share of teachers with teacher education (about two thirds of a standard deviation) increases the probability of high school completion with about half a percentage point. In other words, it takes a 20 percentage point increase in the share of teachers with teacher education to increase the probability of high school completion with one percentage point. In addition, we find that the level of education is not associated with high school dropouts. That is, it does not make any difference whether middle school teachers hold a master's or a bachelor's degree. We discuss that increasing pedagogical competence in the middle schools could contribute to lower high school dropout if the results are due to the pedagogical part of teacher education. However, we also consider that the result could indicate that it is important to work within your field of education.

1. Introduction

Students dropping out of high school is a prominent problem, both for the individual and for the economy as a whole (Belfield & Levin, 2007). Completing high school lowers the probability of receiving public subsidy, increases the participation in the labor market and increases the probability of attaining higher education (Falch & Nyhus 2011; Falch et al., 2010). OECD statistics show that in many countries around 30 percent of students do not complete high school. In Norway, approximately 75 percent of the population under 25 years old had completed high school in 2014 (OECD.org, 2016).

As a step to strengthen the Norwegian school system, the Norwegian government is currently changing the requirements for teacher education. After the policy implementation takes effect in August 2017, teacher education will be a five-year master's program instead of a four-year bachelor degree (regjeringen.no, 2016). Previous research within the field of educational economics mainly focuses on the link between teacher quality and student outcomes in relation to achievement and income, which is in the short and long term. Our work contributes to the literature by examining the effect of teacher quality on students in the medium term, focusing on high school completion. This is the opposite of what we define as a 'high school dropout'; a student who either never starts high school or starts, but never completes high school. Specifically, we assess two key variables. First, if it makes any difference whether middle school teachers have a teacher education which includes pedagogical elements and/or second, if it makes any difference whether middle school teachers hold a master's degree or a bachelor's degree.

In recent years there has been a change in focus amongst researchers from input based to output based teacher quality (Hanushek & Raymond 2001; Hanushek, 2003). Input based research has focused on observable teacher characteristics, such as education, experience, and teachers own academic achievement, while output based teacher quality is based on non-observable characteristics estimating how use of different teacher incentives affect student achievement. Value-added

models measure gains in student achievement; high teacher quality is thereby unobserved, but analyzed through improvement in student achievement, i.e. grades. Empirical findings on input based research have typically found that the relationship between specific teacher attributes and student achievement is weak (Rivkin, Hanushek, & Kain, 2005; Rockoff, 2004). Students that attend better schools are also found to have better short term outcomes measured by examination performances (Jackson, 2010). Output based research has found that students assigned to teachers with high value-added benefit in the long term. They are more likely to attend college and get higher income, live in better neighborhoods, save more for retirement and less likely to have children as teenagers (Chetty, Friedman, & Rockoff, 2011). Students who drop out of high school will be restricted from attaining higher education and thus face limited career possibilities later in life (St.meld. nr.44, 2008-2009). While a value-added model can identify which teacher is effective in improving student outcome, it cannot tell why. The focus has therefore shifted towards classroom strategies and teachers behavior. Since we cannot connect individual teachers to individual students, we are not able to use value-added models to identify how much teachers contribute to student outcomes, using Norwegian data.

We take advantage of rich administrative panel data, where the data on teacher characteristics has not yet been used. Middle school students are of interest, as school up to this point has been mandatory and therefore students are not yet affected by their own educational choices. Regarding the students, individual characteristics, in particular grades from middle schools, are related to whether students complete high school (Byrhagen, Falch, & Strøm, 2006). Motivation and expectations about the rewards from graduation are also found to be important when explaining high school completion (Eckstein & Wolpoin, 1999). They use the first registered grades from compulsory schooling, namely the 1986 cohort, who completed middle school in 2002. Furthermore, June C. Rivers and William L. Sanders (2002) find that at least four years after students have left a teacher's tutelage, the effect of that teacher on the student is still measurable. Subsequently, we believe that we will be able to capture a relationship between middle school teachers education and high school completion.

Quasi-experiments are regularly used to study the causal effect of high school dropouts. One relevant example is the quasi-experimental study on Norwegian data from middle school that found no effect of class size on achievement, independent of teacher characteristics (Leuven, Oosterbeek, & Rønning, 2008). However, they found indications that the students from classes with teachers with a lower level of education benefit more if the classes are small, than students from classes with more educated teachers. A useful and common starting point is to analyze correlations and estimate regression models in administrative data. This is a way of understanding the relationship between middle school teacher education and high school completion when treatment is not randomly assigned.

We exploit within middle school variation in teachers' education over time. We use students from eleven different birth cohorts to examine the relationship between teacher education and the probability that a student completes high school on time. The effect of teacher quality can be difficult to identify with administrative data, due to the complexity of the relationship and that student outcome is affected by many other factors than quality of teachers. We solve this by controlling for middle school characteristics, student characteristics, constant middle school factors, and variation between birth cohorts. Since the Coleman report from over fifty years ago, educational economists have tried to figure out what makes a good teacher, and separate these effects from family characteristics in order to measure teacher effect on student outcomes (Coleman, 1968; Downey & Condron, 2016). They, among others, have tried to isolate the direct relationship between student characteristics and teacher quality from the effect of neighborhoods or schools (Downey & Condron 2016; Rivkin, Hanushek, & Kain, 2005). In Norway, as for many other countries, student background, such as parent characteristics and home location, matters for their achievement (Statistics Norway, 2017). Hægeland, Raaum, and Salvanes (2005) find evidence of compensating resource allocation, meaning that the parents compensate by spending resources on their children if they are not satisfied with the teachers or the school. They also find that there is geographical sorting amongst families and teachers in Norway. The parents´ decision on where to settle down, does to some extent, depend on school characteristics such as teacher quality (Nechyba, 2006;

Fiva & Kirkebøen, 2008; Machin & Salvanes, 2016). Several empirical studies show that teachers tend to choose where to work based on student characteristics such as minorities and achievement (Hanushek, Kain, & Rivkin, 2004; Jackson, 2009).

The literature on observable teacher characteristics covers teachers own achievement, experience and education among other characteristics. For example, for student achievement, teacher experience is found to matter most the first three years in the profession (Wiswall, 2013). Though, experienced teachers tend to choose where to work more systematically (Hanushek, Kain, & Rivkin, 2004). Several scholars provide evidence of a positive influence of teachers' academic achievement on student achievement (Hanushek, Piopiunik, & Wiederhold, 2014; Hanushek & Rivkin, 2006; Clotfelter, Ladd, & Vigdor, 2006 and 2007). For input based research regarding teacher education, only a small share has found a master´s degree to be positive and significant for student performance (Hanushek, 1997). U.S. studies have also focused on salary as a driver for teacher distribution (Hanushek & Rivkin, 2006). Regarding teacher salary in Norway, the variation across schools and locations is expected to be small and much smaller than those observed in the U.S., due to the collective centralized agreements. Salary is also found to be more important for the decision of whether to become a teacher than whether to work at a specific school or even quit teaching (Hanushek & Rivkin, 2006). Teacher sorting may reflect several factors such as personal preferences and the labor market situation amongst other examples.

We find that there is a positive relationship between middle school teachers with teacher education and high school completion. Specifically, we find that a 10 percentage point increase in the share of teachers with teacher education (about two thirds of a standard deviation) increases the probability of high school completion with about half a percentage point. In other words, it takes a 20 percentage point increase in the share of teachers with teacher education to increase the probability of high school completion with one percentage point. In addition, we find that the level of education is not associated with high school dropouts.

The thesis is structured in the following order. In the next section, an overview of the Norwegian school system and institutional settings is given. In Section 3 we describe the data we use and present descriptive statistics. In section 4, the empirical strategy is explained and the properties of the main model with middle school fixed effects is identified. In section 5 we present the results from our analysis. Lastly, in section 6, we discuss and conclude upon the results.

2. Institutional Settings

2.1 School System

In this thesis, we study students in the birth cohorts 1967-1977 who attended middle school in 1981-1991, which we match with a rich administrative data set on teacher characteristics. As can be seen in table 1, for students of these birth cohorts, compulsory school consists of nine years: grades 1–6 in primary school and grades 7–9 in middle school. In Norway, the municipalities are responsible for allocating resources to the middle schools. There are few private schools, and they are either religious or have alternative pedagogics.

Table 1: The Norwegian school system

Age	Grade	Educational level	
7 to 12	1 to 6	Primary school	Compulsory education
13 to 15	7 to 9	Middle school	
16 to -	1 to 3	High school	

Parents do not choose middle school for their kids per se, as each school has its own catchment area. Parents can apply for transfer of their child to another school, however whether the transfer is approved depends on subjective characteristics and availability at the school. A middle school can be mixed with a primary school. If not, the primary school often has one specific middle school

where the students are enrolled. Mixed primary and private schools are excluded from this thesis. In the relevant time period, children enrolled into primary school in August the year they turned seven years and middle school the year they turned thirteen. The program for international student assessment (PISA) was not introduced in Norway before 2003, national tests were introduced in 2007 and grades from middle school completion were first registered by SSB in 2002.

Students are assigned their middle school teachers in the 7^{th} grade and the combination of teachers is normally relatively consistent for all three years at middle school. Hence, a new teacher is usually assigned to the 7^{th} grade. Therefore, we connect students with the average teacher education in the first year they attend middle school.

After completing middle school, students can proceed directly to high school the year they turn 16. There are two main specialization alternatives at high school, that is, general and vocational programs. The educational length is three years for the general studies and four years for the vocational programs. For the latter, typically two years in school followed by two years with trainee practice for a company. Before the reform of 1994, students were not guaranteed to get into a high school. Accordingly, our results may differ as all students today have the right to attend high school.

There are differences in quality between Norwegian schools. For example, Fiva and Kirkeøen (2011) found that moving a student from a school in the bottom 25 percent to a school in the top 25 percent would improve student's grade-point average. There are also found to be big difference in resource spending , and the variation is especially prominent between middle schools (Borge & Naper, 2005). These differences among the middle schools may be important for high school completion.

2.2 Teacher Education

In the 1980´s, teacher education could be either two or four years depending on high school specialization. There were also several ways to become a teacher.

Through teacher education one could become a regular teacher, a subject teacher or a vocational teacher. Individuals with general higher education could also become teachers by taking an additional year of pedagogics, though many worked as a teacher without a teacher education. Today, all first to tenth grade teachers take a four-year bachelor's degree. With the new requirements, all new teachers must take a master's degree (regjeringen.no, 2016). The teacher salaries are determined through collective agreements.

For higher education in general, the terms master's and bachelor's degrees were not used until 2003. We will use the terms "short higher education" and "long higher education". Short higher education includes all programs with less than five years of higher education. Therefore, short higher education is similar to today´s bachelor's degree, consisting of three years at university or university colleges. Long higher education corresponds to a master's degree or more.

3. Data and Descriptives

We use four administrative data sets. Because the data sets are administrative, we face only slightly or no measurement error in the dependent and independent variables, such as teacher education codes and whether and when students have completed high school. We combine data regarding teachers, which we aggregate at middle school level, with information of the students at individual level.

3.1 Teacher register

The teacher register is a panel data set which is an ended series of administrative data with information about teachers on an individual level that has never been used before. We use information regarding the teachers education from this register covering all schools in Norway, updated on a yearly basis over the years 1981 to 1991. We thereby limit our focus to these years. The data set covers characteristics on the individual level of each teacher. We use education codes, age, which school level they teach at, professional title, the teachers' home municipality, service condition and school identification. We restrict the data to

those who teach at middle school level. These observations cover all employees that can affect the teaching, such as principals, inspectors and teachers. At small schools, principals job description often includes teaching. This is normally not the case at bigger schools, meaning the effect of a teaching principal is very small. In addition to the different types of teacher education, teachers hold a variety of educational specialisms, from economic and administrative to vocational backgrounds.

Because the data set is aggregated at middle school level and year, each middle school's teacher education is measured by its teachers' average education in a certain year. This way we connect birth cohorts of the students to the aggregated teacher education at the middle school that the students attend. As described in chapter 2, a certain composition of teachers typically follows the same cohort for all three years. The regressor of interest is therefore 7^{th} grade middle school teacher's average education. In the data, there is variation in middle schools teacher education over time, which this is important for the estimation of our main model.

Two different key educational measures on teachers will be used as regressors, separating teacher education in two different ways. The quality measures are long higher education and teacher education. The first educational measure is based on the separation between teacher education and other educations from university or university-college. Hence, the share of teachers with teacher education as a quality measurement includes all levels of teacher education, such as subject teachers, vocational teachers, ordinary teachers and those with general higher education and additional pedagogics. The remaining share includes all teachers with other types of educations. The second educational measure is separating long higher education from short higher education. As described in chapter 2, short higher education includes all programs with less than five years of higher education, including short higher teacher education. Long higher education is defined as teachers with minimum five years of higher education independent of program specification, including long higher teacher education.

3.2 Register of the population's level of education

The register of the population's level of education is a panel data set which covers the population's education on an individual level; both ongoing and attained levels, updated since 1970. We use information regarding students' high school completion and what type of specialization was chosen. Education data tracks people's education from upper secondary education to PhD-level. We define "high school completion on time" as completion within three years if the student attended a general study program and completion within four years if the student attended a vocational program. The high school dropouts are therefore a sum of those who did not start high school and those who started but did not complete.

3.3 Register of compulsory Schooling

The register of compulsory schooling is a panel data set which covers information about all schools on a class level. Information on middle school characteristics is collected from the compulsory schooling information system (GSI). This data set contains information of all compulsory schools in Norway since 1970. We use information regarding class size and minorities to construct control variables.

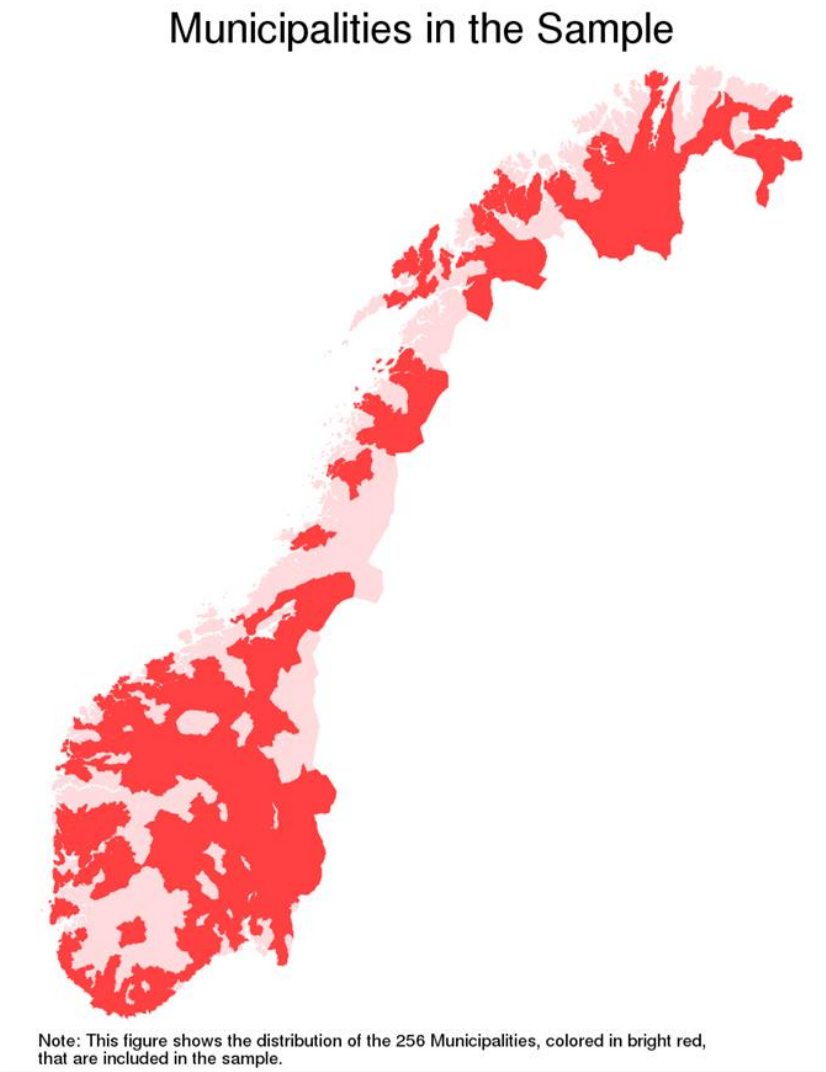
3.4 Middle school register

The middle school register is a cross sectional data set which covers information about students at an individual level at their last year at middle school. The data set contains the name and location of which middle school the students attended. From this data set, we also get the students' parents' education, home municipality and the middle school graduation year. Only students who completed middle school on time are included. The data on year of middle school completion is shown per school semester. The data sets we use give us no connection between individual students and primary or high school.

To summarize, our sample contains information about teachers between the years 1981-1991. While there exists records from about 450 municipalities in Norway, we choose to include the municipalities that have not merged and that have students from the cohorts between 1967-1977. In more than one third of the

municipalities, the schools are either mixed primary and middle schools or there is no middle school, indicating that students go to a middle school in a neighboring municipality. The sample therefore covers 256 municipalities (Fig.1). If we were to consider mixed primary and middle schools we would get a less precise result, as students from mixed schools have not had tutelage of all middle school teachers. Therefore, only middle schools are considered in this thesis. The excluded municipalities are located in the inland, mid, and northern Norway. In total, there are 452 middle schools in the sample.

Fig. 1



3.4 Descriptive Statistics

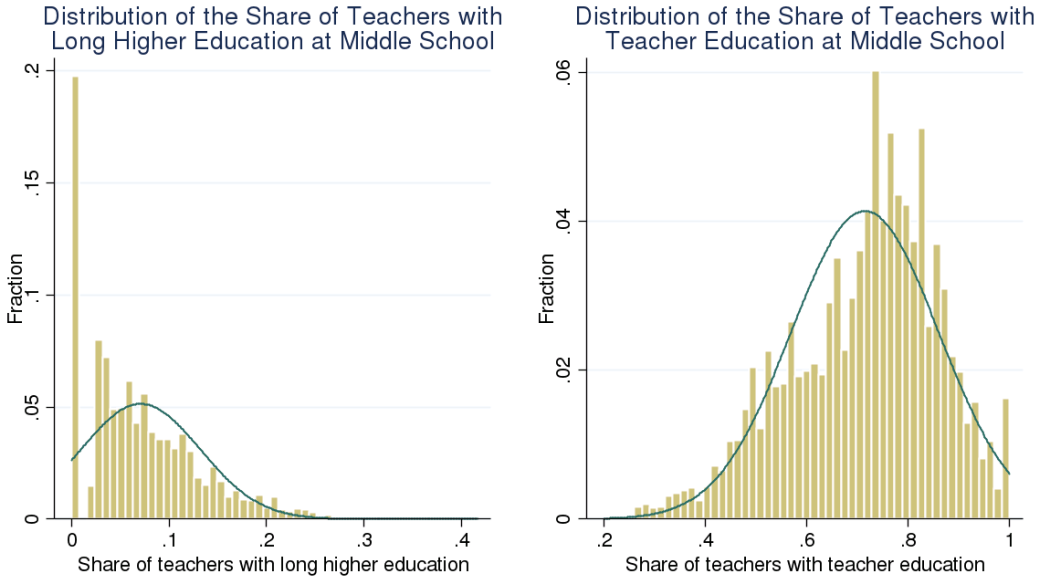
Table 2 shows descriptive statistics of the sample. The share of teachers with long higher education is about 8 percent, while the share of teachers with teacher education is 70 percent (Fig. 2). None of the middle school teachers have high school as their highest achieved education. The share of middle school teachers who are working full time is 77 percent and there is almost as many female as male teachers. At the middle schools the share of teachers with long higher education varies between 0 and approximately 30 percent. The share of teachers with teacher education varies between 20 to 100 percent.

Table 2 Summary Statistics

Variable	N	Mean	Std. Dev.
Middle School characteristics			
Class size*	370140	106.2	39.92
No. of minorities	371539	0.43	2.25
Middle school Teacher characteristics			
Share with long high education	371539	0.08	0.06
Share with teacher education	371539	0.70	0.15
Average age	371539	41.22	3.30
Average age^2	371539	1790	281
Share working full time	371539	0.77	0.11
Student characteristics			
High school on time (Dummy)	371539	0.47	0.50
High school within five years (Dummy)	371539	0.62	0.48
High school achieved within 2015 (Dummy)	371539	0.80	0.40
Age when completed high school	279066	20.78	4.12
Students highest education**	371539	4.72	1.62
Fathers with higher education (Dummy)	371539	0.22	0.41
Mothers with higher education (Dummy)	371539	0.17	0.37
Fathers highest education***	368990	3.56	1.52
Mothers highest education ***	368954	3.25	1.37
Municipalities	256		
Middle schools	452		
Students	371539		

Notes:
Summary statistics on the sample for the cohorts 1967-1977,
middle school characteristics and teacher characteristics are aggregated at middle school and cohort levels,
while student characteristics are observed at individual level.
* total class size per cohort and year in 7th grade.
**NUS 2000 codes. The number of the code reflects the completed level of education.
2= middle school; 3= ongoing high school; 4=completed high school; 5=additional courses at high school level;
6=short higher education; 7=long higher education; 8=doctoral degrees.
*** short higher education or above.

Fig. 2

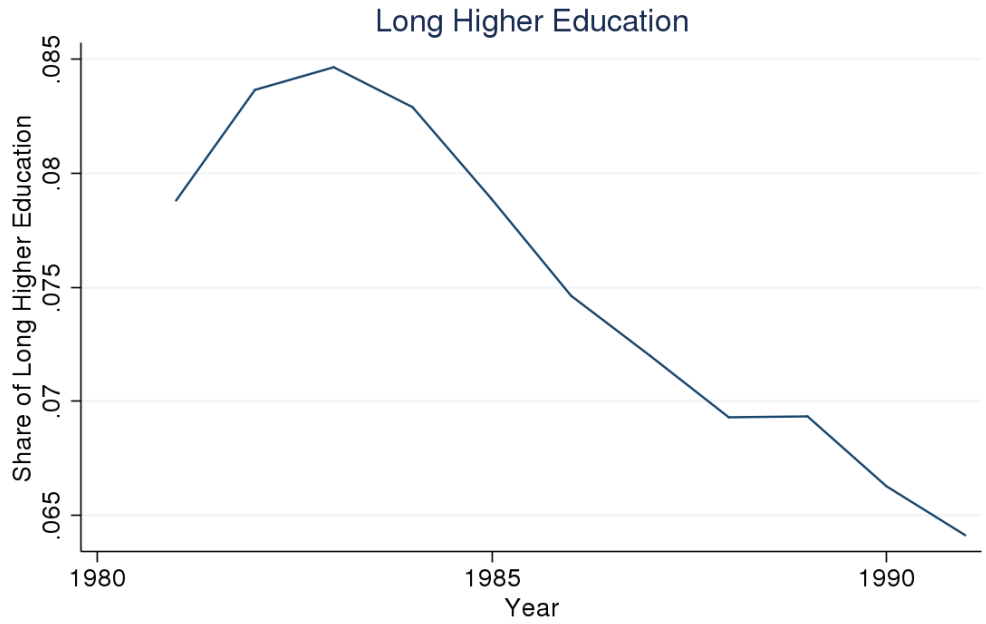


Note: The bins in the left graph shows the fraction of middle schools share of teachers with long higher education, Y-axis, by the average share of teachers at the 452 middle schools in the sample, between 1981-1991 ,X-axis. The green line shows the normal distribution of the share of teachers with long higher education.
The bins in the right graph shows the fraction of middle schools share of teachers with teacher education (Y-axis), by the average share of teachers at the 452 middle schools in the sample, between 1981-1991 (X-axis). The green line shows the normal distribution of the share of teachers with teacher education.

There is an increase in the share of teachers with teacher education in the time period, while the share of teachers with long higher education is decreasing (Fig 3a and 3b). Hence, we see a trend that the teacher profession is being more formalized and a carrier in itself. Furthermore, for the analyzed time period we observe teachers in the profession for on average 8.5 out of 11 years. We observe within middle school variation in the teacher's education over time. This variation could be due to teachers moving, having paternity or maternity leave, changing profession, retiring etc.

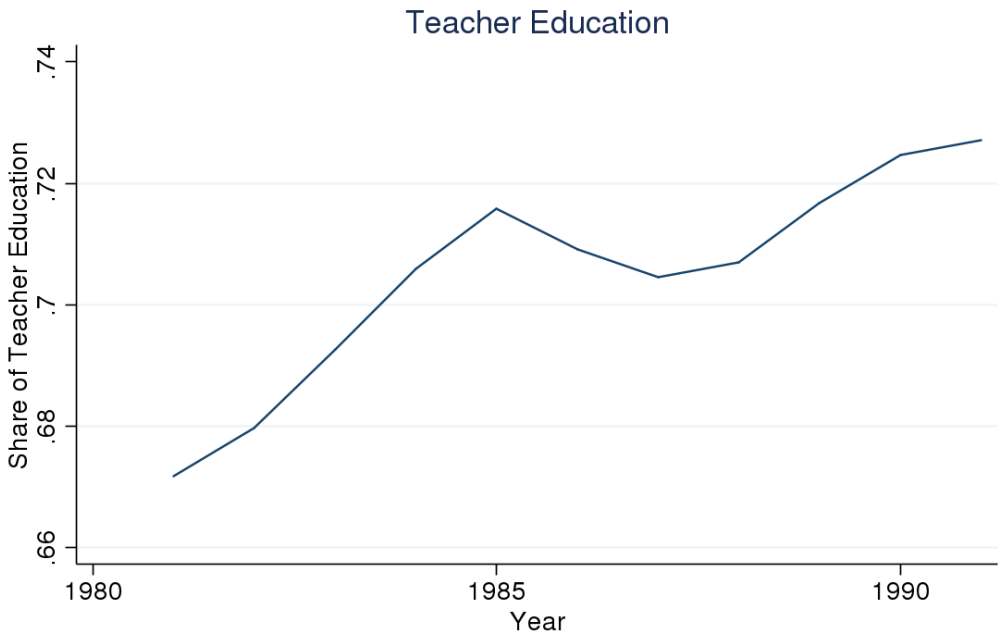
Just below 45 percent of the students complete high school on time (Table 2). Within the time limit of the data (year 2015), just under 80 percent of the students ever complete high school, which illustrates that the drop out ratio is large. Of those who complete middle school on time, about 85 percentage start directly at high school and 3 percent never start high school. As can be seen in Fig. 4 of the high school completion ratio per cohort, there is a trend in the data that students of later cohorts to a larger extent complete high school. Furthermore, some middle schools have students with lower high school completion overall years, indicating that there could be variation in school quality (Fig. 5).

Fig. 3a



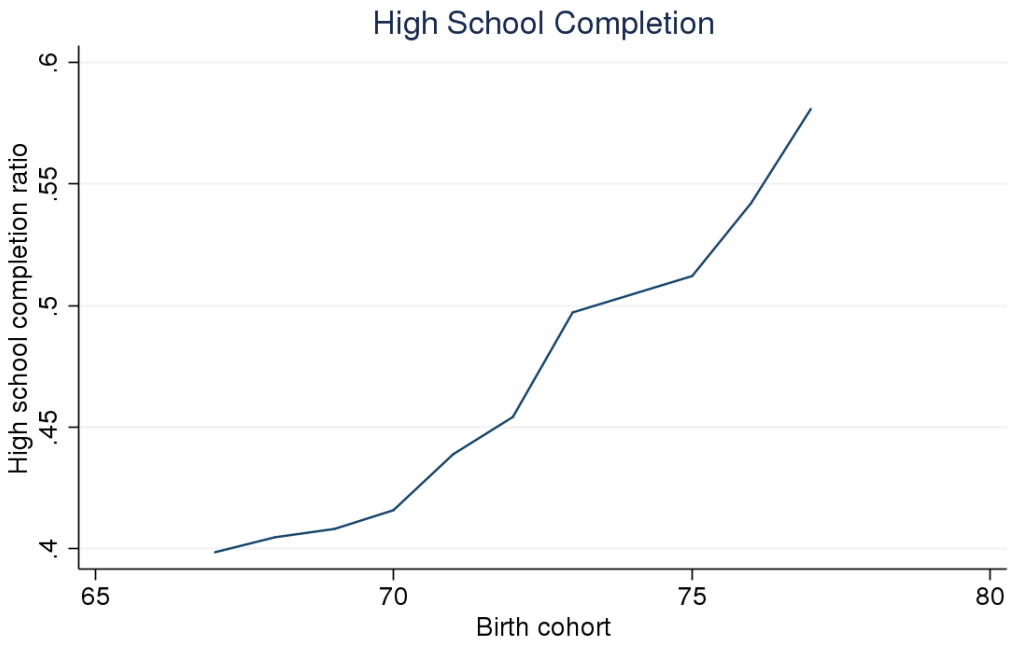
Note: This graph show the development over time in the share of teachers with long higher education. We use the sample of all teachers aggregated, at 452 middle schools that run all the years 1981-1991.

Fig. 3b



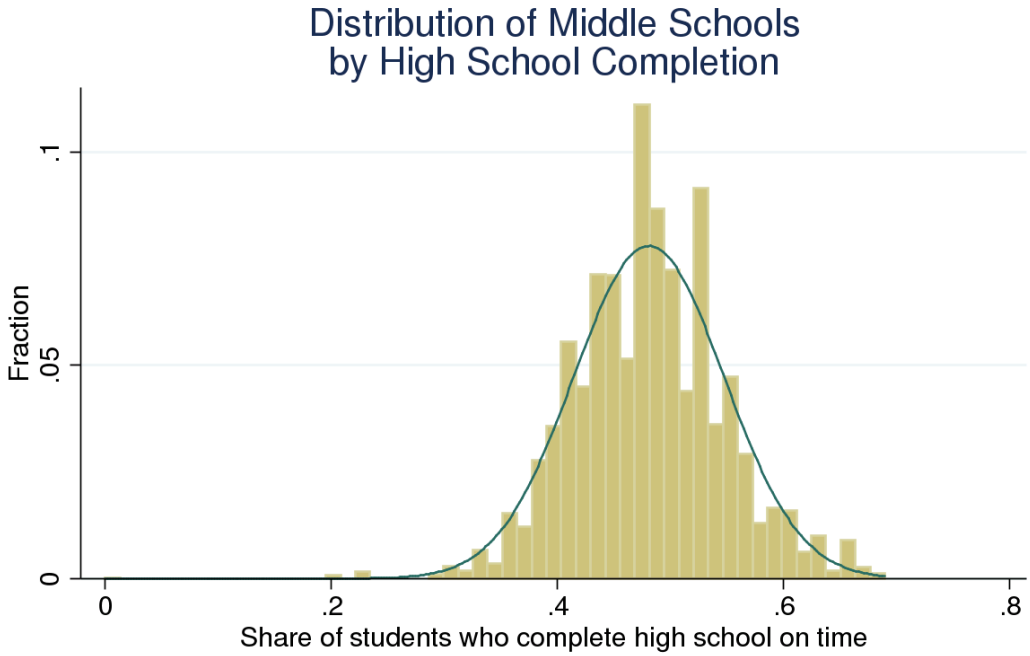
Note: This graph show the development over time in the share of teachers with teacher education. We use the sample of all the teachers, aggregated at 452 middle schools that run all the years 1981-1991.

Fig. 4



Note: This graph show the the share of students who complete high school on time, by their birth cohort. The sample with 371,539 students attaining the 452 middle schools that run between 1981-1991, except cohort 1974 that lack many registrations.

Fig. 5



Note: The bins in this figure shows the fraction of the 452 middle schools by the share of studenst who complete high school on time. The green line shows the normal distribution of the share of students who complete high school on time.

The students' average highest achieved level of education is above high school, but below the level of education that corresponds to additional courses at high school level. While the share of mothers and fathers of the students who have completed higher education is 17 and 22 percent respectively, the parents have on average started high school but not completed, see Table 2. The development in mothers and fathers share of higher education increase over the years (A1).

High school completion is positively correlated with teachers with long higher education and negatively correlated with teachers with teacher education, see A2. As expected, high school completion is also positively correlated with parents education. Teachers with teacher education correlates negatively with parents education and teachers with long higher education correlate positively with parents education. Long higher education for teachers and high school on time correlates positively with the share of minorities at the middle school. The correlation between the share of teachers with long higher education and the share of teachers working fulltime is negative.

We use maps that separate the different municipalities to analyze the distribution of teacher and student characteristics in the data, see A5-A10. The distribution of teachers with teacher education is especially prominent in the municipalities in the southern parts of Norway (A5). In addition, some scattered municipalities have a larger share of teachers with teacher education. The share of teachers with long higher education is lower in the mid of Norway compared to the south and north (A6).

The high school completion ratio on time is lower in the northern parts of Norway and in the inland (A7). Additionally, as can be seen in table 2, while the average age of completing high school is about 21 years, the students in the northern and mid-Norway are older than students in the east, western parts and the big cities when completing high school (A8).

The geographical distributions have patterns that to some extent are repeated in several of the variables. Both teachers with long higher education, parents with higher education and students who complete high school on time are centered in

and around the major cities. The pattern visualizes spatial sorting of high ability parents and high ability teachers that is present between municipalities. It is hard to control for all these factors. Therefore if we were to estimate a cross section OLS we expect it to be biased. To address this problem we take advantage of the within middle school variation over time.

4. Empirical Strategy

We estimate a linear probability model with middle school fixed effect for two explanatory key variables. Our outcome variable, y_{ics}, is an indicator function that takes the value 1 if student *i* complete high school within the stipulated time, and 0 if not. Whether student *i* from birth cohort *c* at school *s* complete high school or not, y_{ics}, is described by the following equation:

$$(1) \qquad y_{ics} = \beta_0 + \beta_1 T_{cs} + \delta w'_{cs} + \varphi x'_i + \gamma_s + \theta_c + \mu_{ics}$$

Where teacher education, T_{cs}, is either the share of teachers with long higher education or the share of teachers with teacher specific education, at middle school *s*, for students of birth cohort c. The parameter of interest, β_1, captures the effect of teacher education on the probability that student *i* complete high school, holding all other variables constant. The vector w'_{cs} consists of other observable middle school characteristics, while x'_i is a vector of observable student characteristics. w'_{cs} includes the share of full time teachers, average age of teachers and age squared, the number of minorities per cohort and year, and total class size per cohort, while x'_i includes each students mothers and fathers education level. γ_s captures all middle school factors that do not vary over time, such as prejudice, resource structure, unobservable teacher quality, and location. Students of a certain birth cohort may differ in many aspects relative to other birth cohorts and therefore a birth cohort fixed effect, θ_c, is included to control for this variation, which also can be seen from Fig. 3. For example, if certain birth cohorts have been exposed to a national school reform that other birth cohorts are not exposed to, then this will be captured by θ_c. The error term, μ_{ics}, denotes any

measurement error and all variation in y_{ics} that is not explained by the other variables. We cluster the standard errors at middle school level to take intra-class correlation into account.

As we use panel data we observe the same middle school over time. With middle school fixed effect we extract the average from the middle schools so that what is left is the within-school variation over time. This way variation across the middle schools is not used when estimating the regression coefficients. We thereby take advantage of the fact that we observe variation in the teacher education composition at a certain middle school over several student cohorts. Hence, we avert many sources to bias that would otherwise affect the estimated results.

If the selected control variables, discussed below, pick up all relevant differences between students and teachers, the causal effect of teacher education is isolated by the conditional independence assumption (Angrist & Pische, 2009). Middle school characteristics are expected to introduce bias in the explanatory variables if we do not control for factors that give rise to sorting behavior of different types of teachers and families (Wooldridge, 2010). We therefore control for several middle school characteristics that could otherwise lead to omitted variable bias (OVB), that occur since several middle school characteristics are correlated with both teachers type of education and student high school completion. Student composition matters for the characteristics of a middle school. We include minorities as control variable as it might have an effect on the teacher-sorting problem (Falch & Ronning, 2007). Teacher age and age squared are included as control variables as a compliment to experience. The quadratic term is included, as the effect of longer experience is believed to be positive but decreasing with time. Whether the teachers are working full time or not is a characteristic that may influence to which extent a certain teacher affects the students and is included as control variable. Individual student characteristics, for example high achieving students, may also to a larger extent complete high school. However, we do not have data on student achievement such as grades. Not controlling for achievement could introduce positive bias in the results.

To isolate the causal effect of teacher education on high school completion, we rely on the conditional mean independent assumption:

$$E(\mu_{ics}|T_{cs},\ w'_{cs}, x'_i, \gamma_s, \theta_c) = E(\mu_{ics}|\ w'_{cs}, x'_i, \gamma_s, \theta_c)$$

Hence, once we control for middle school characteristics (w'_{cs}), student characteristics (x'_i), constant and unobserved middle school factors (γ_s) and variation between birth cohorts (θ_c), then T_{cs} can be as good as randomly assigned. If the conditional mean independence assumption holds, $\widehat{\beta_1}$ has a causal interpretation.

It is found that teachers tend to prefer to work close to where they are born or places similar (Boyd et al., 2005). If teachers prefer to move to their home district then the school characteristics might be less important for the "choice of workplace". This is to some extent believed to be applicable in Norway, a geographically spread country with many disperse rural areas. Therefore, geographical teacher sorting is expected to be smaller than it would be if home municipality did not matter. Nevertheless, we still expect teacher and families to sort within neighborhoods, particularly in bigger municipalities where schools are located in areas with different status and within relative short distance of each other. The problem that middle schools differ in constant unobservable characteristics, can affect the sorting of both teachers and families as these between-school differences can make some middle schools more or less attractive than other middle schools.

The strength of the model is that it relies on within middle school variation. With middle school fixed effect we avoid neighborhood sorting, typically along the socioeconomic dimension, since we control for fixed omitted variables that are time invariant (Angrist & Pischke, 2009). Middle school fixed effect will also correct for the same geographical sorting as the municipality fixed effect. We observe that there is variation in teacher education within schools over the years 1981-1991, as described in section 3.4. For a certain cohort of students, a certain composition of teacher education is therefore (argued to be) as good as random.

We argued earlier that when a new teacher enters a middle school, this teacher is assigned to a new class of students and therefore do not affect the older students. However, this argumentation may not hold at a small middle school where all the teachers may alter between all the classes, which can lead to downward bias. Moreover, a new teacher may affect the older students indirectly, through cooperation with colleagues. Hence, there might be teacher peer effects (Jackson & Bruegmann, 2009).

5. Results

The estimated results for model (1), with middle school fixed effect are presented in Table 3, where the first three columns refer to the key variable share of teachers with teacher education, and columns four to six refer to the key variable share of teachers with long higher education.

We find that a 10 percentage point increase in the share of teacher with teacher education (about two thirds of a standard deviation) increases the probability of high school completion with about half a percentage point. In other words, it takes a 20 percentage point increase in the share of teachers with teacher education to increase the probability of high school completion with one percentage point. In comparison to the large variation in high school completion rates across middle schools, ranging from about 0.2 to 0.7 (see Figure 5), we consider the effect to be relatively small in magnitude. Though, we find it to be an interesting result as literature typically do not find a clear relation between observable teacher quality and student outcomes (Jepsen, 2005; Hanushek & Rivkin, 2006; Rivkin, Hanushek, & Kain, 2005; Rockoff, 2004).

Table 3: Regression Results

Dependent variable: High school completion							
Middle School Fixed Effect							
	(1)	(2)	(3)		(4)	(5)	(6)
Share of Middle School Teachers with Teacher education	0.0515**	0.0513**	**0.0438***	Share of Middle School Teachers with Long higher education	-0.0740	-0.0068	-0.0184
	(0.02)	(0.02)	(0.02)		(0.04)	(0.04)	(0.04)
Fathers education		0.1875***	0.1874***	Fathers education		0.1875***	0.1874***
		(0.00)	(0.00)			(0.00)	(0.00)
Mothers education		0.1123***	0.1124***	Mothers education		0.1123***	0.1124***
		(0.00)	(0.00)			(0.00)	(0.00)
Minorities			-0.0002	Minorities			-0.0002
			(0.00)				(0.00)
Full time			-0.0034	Full time			-0.0037
			(0.02)				(0.01)
Teachers age			0.0158***	Teachers age			0.0170***
			(0.01)				(0.01)
Teachers age^2			-0.0002**	Teachers age^2			-0.0002***
			(0.00)				(0.00)
Class size			0.0001**	Class size			0.0002**
			(0.00)				(0.00)
Adjusted R-Squared	0.0688	0.0688	0.0688	Adjusted R-Squared	0.0687	0.0687	0.0688
N	371539	371539	371539	N	371539	371539	371539

Notes:

Significance levels * p < .1; ** p< .05; *** p < .01. Estimated results of the main model (1) from empirical strategy.

Middle school clustered standard errors in parenthesis. All regressions control for birth cohort of the students (dummies).

Mothers and fathers share with higher education are dummies on individual level of the students, while all other control variables are the share on school and cohort levels.

The estimated results, with and without controlling for parents education are similar and significant at five percent level, see column (1) and (2). When including all control variables, $\widehat{\beta_1}$ decreases a little and is significant at ten percent level. The fact that the estimated results in columns (1) to (3) estimates rather similar results, with and without control variables, strengthens the validity of the model and the sign of sorting within the variables is as good as absent. The adjusted R-square is low and stable. As expected and in line with the literature, we find that the coefficients of both mothers and fathers education are large and significant at one percent level. This shows that parents education to a large extent explain whether the students complete high school. The middle schools average

teacher age has a positive significant effect and age squared has a significant negative, but close to zero effect. This is in accordance with empirical result which finds that teacher experience matters most the first few years (Wiswall, 2013). The size of the birth cohort per middle school (class size) has an effect that is close to zero, though significant. This is a highly discussed topic, however our result seems reasonable as most researchers find that class size has no or little effect on student achievement (Leuven, Oosterbeek, & Rønning, 2008; Rivkin, Hanushek, & Kain, 2005). The effect of minorities is close to zero, negative and not significant. However, this variable is aggregated to school and cohort levels, and in contrast to the other variables to some extent incompletely reported. Preferably we would have used the national immigrant register, and then the information would be at individual level. Though, in the 1980's the share of immigrants was much lower than to day. In addition, the effect of the share of teachers working full time is also negative and not significant.

As seen in Table 3, column (4)-(6), with middle school fixed effect, for the regression of long higher education on the completion at high school, neither of the estimated results are significant. For all columns, (4)-(6), the model estimates a β_1 which is negative and insignificant. The estimated result decreases for each column, when gradually including control variables. Hence, we find no relationship between the level of education and high school completion.

The results in table 3 differ from the results when estimating the model both for the naïve OLS and with municipality fixed effects, for teacher education on high school completion, see Table A3. The estimated results for the naïve OLS are all close to zero and insignificant. While the estimated results with municipality fixed effects are small, decreasing, changing from positive to negative, and insignificant. This gives indication of between-municipality sorting. Since the model improves when regressing with middle school fixed effect the estimated result for the municipality fixed effect are biased due to within-municipality sorting.

Signs of sorting are especially clear when we regress long higher education on high school completion. The estimated results in the naïve OLS and with

municipality fixed effects, for long higher education on high school completion, changes when we gradually include the control variables, due to geographical sorting between municipalities of both families and teachers (Table A4). These results differ a lot from the model with middle school fixed effects. Hence, the effect of long higher education on high school completion is overestimated in the naïve OLS and with municipality fixed effects.

One could further check the robustness of the results by excluding teachers who switch schools to make sure that they do not drive the results, and investigate whether the observable characteristics differ between teachers with different educations.

6. Discussion and Conclusion

In this thesis, we investigate whether middle school teachers' education is a determining factor in the probability that students complete high school, as middle school factors has been shown to be one of the prominent explanations for high school completion (Byrhagen, Falch, & Strøm, 2006). We focus on two key variables at middle school level, first teachers with and without teacher education, and second, teachers with short and long higher education. We combine four data sets to estimate our main model. The detailed data on teacher characteristics has been of particular interest, as it has not been used before. The model has advantages as it uses within middle school variation in teacher education over time. This reduces omitted variable bias caused by teacher and family sorting, both cross sectional and within municipalities and neighborhoods, and subsequently aims to capture the causal relationship of interest.

We find that there is a positive relationship between middle school teachers with teacher education and high school completion. Specifically, we find that a 10 percentage point increase in the share of teacher with teacher education (about two thirds of a standard deviation) increases the probability of high school completion with about half a percentage point. In other words, it takes a 20 percentage point increase in the share of teachers with teacher education to

increase the probability of high school completion with one percentage point. In comparison to the large variation in high school completion rates across middle schools, ranging from about 0.2 to 0.7 (see Figure 5), we consider the effect to be relatively small in magnitude. Though, we find it to be an interesting result as literature typically do not find a clear relation between observable teacher quality and student outcomes (Jepsen, 2005; Hanushek & Rivkin, 2006; Rivkin, Hanushek, & Kain, 2005; Rockoff, 2004).

In addition, we find that the level of education has no relation with high school dropouts. That is, whether middle school teachers hold a master's or bachelor's degree does not make any difference on the probability that a student completes high school. Moreover, in line with both previous research and our expectations, we find that what best explains the high school completion is the parents' level of education.

Consequently, increasing pedagogical competence in the middle schools could contribute to lower high school dropout if the results are due to the pedagogical part of teacher education. Then the result could indicate that, at middle school level, pedagogical knowledge is more important than subject specific knowledge. However, the result could also indicate that it is important to work within your field of education. For example, the result might capture that teachers with teacher education are more motivated than teachers with another education, because the latter work in another profession than they are educated within.

This should be relevant information for policy makers aiming for evidence-based policies. The new requirement for teacher education combines long higher education with pedagogic education. If the results are caused by pedagogical knowledge, then master educated teachers after the policy implementation might positively affect the student's outcome, here measured as completing high school. However, if the results are caused by mismatching between working profession and education, the results might indicate that one should rather focus on attracting the "right" people to the teacher education.

In any case, the effect of the new teacher requirements can only be assessed through further research. As the teacher quality is hard to capture by observable data we believe that classroom studies are needed to further investigate what makes a good teacher. This can be done by evaluating teacher classroom strategies.

These days, the government facilitates for randomized experiments when proposing new requirements. By doing this, researchers can capture the causal effect and determine whether the requirements are reasonable. A good example where the government have tried to do an experiment is the project “ny giv” or in English “new motivation (Huitfeldt, Kirkebøen & Rønning, 2016). However, it is hard to carry out an experiment as many steps can easily fail, in both the design and the implementation of the experiment. Though, we believe in smaller experiments such as “program for better high school completion” where researchers have more control when implementing the experiments (regjeringen.no, 2016).

References

Angrist, J. D., & Pischke, J. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

Belfield, C. R., & Levin, H. M. (2007). *The economic losses from high school dropouts in California*: University of California, Santa Barbara.

Borge, L. & Naper, L. R. (2005). Ressurssituasjonen i grunnskolen 2002-2004. *Senter for Økonomisk Forskning,Rapport nr. 06/05.*

Byrhagen, K. N., Falch, T., & Strøm, B. (2006). Frafall i videregående opplæring: Betydningen av grunnskolekarakterer, studieretninger og fylke. *Senter for Økonomisk Forskning,Rapport nr. 08/06.*

Chetty, R., Friedman J. N., & Rockoff, J. E. (2011). *The Long-Term Impacts of Teachers Teacher Value-Added and Student Outcomes in Adulthood*. Cambridge, Mass: National Bureau of Economic Research.

Coleman, J. S. (1968). The concept of equality of educational opportunity. *Harvard Educational Review*, 38 (1), 77-22. DOI: 10.17763/haer.38.1.m3770776577415m2?code=hepg-site

Cotfelder, C.T., Ladd, H.F., and Vigdor, J.L. (2006). Teacher-Student Matching and the Assessment of Teacher Effectiveness. *Journal of Human Resources* 41 (4), 778–820. DOI: 10.3368/jhr.XLI.4.778

Clotfelter, C.T., Ladd, H.F., and Vigdor, J.L. (2007). How and Why Do Teacher Credentials Matter for Student Achievement? *Economics of Education Review* 26 (6), 673-682.

Downey, D. B., & Condron, D. J. (2016). Fifty Years since the Coleman Report: Rethinking the Relationship between Schools and Inequality. *Sociology of Education*, 89 (3): 207-220. doi: 10.1177/0038040716651676.

Eckstein, Z. & Wolpin, K. I. (1999). Why youths drop out of high school: The impact of preferences, opportunities, and abilities. *Econometrica*, 67 (6), 1295-1339. DOI: 10.1111/1468-0262.00081

Falch, T., Borge, L., Lujala, P., Nyhus, O. H., & Strøm, B. (2010). Årsaker til og konsekvenser av manglende fullføring av videregående opplæring. *Senter for Økonomisk Forskning, Rapport*, 3 (10).

Falch, T., & Nyhus, O. H. (2011). Betydningen av fullført videregående opplæring for sysselsetting og inaktivitet blant unge voksne. Søkelys på arbeidslivet. *Senter for Økonomisk Forskning* 28 (04): 285-301.

Fiva, J., & Kirkebøen, L. (2011). Information Shocks and the Dynamics of the Housing Market. *The Scandinavian Journal of Economics*, 113(3), 525-552. DOI: 10.1111/j.1467-9442.2011.0165.x

Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. In *Educational Evaluation and Policy Analysis,*19 (2), 141-164.

Hanushek, E. A., & Raymond, M. E. (2001). The confusing world of educational accountability. *National Tax Journal*, 54 (2): 365-384.

Hanushek, E. A. (2003). The Failure of Input-based Schooling Policies.(Features: Education)(Author abstract). *Economic Journal*, 113 (485): F64-F98. DOI: 10.1111/1468-0297.00099.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Why Public Schools Lose Teachers. *Journal of Human Resources*, 39 (2): 326-354.

Hanushek, E., Piopiunik M., and Wiederhold, S. (2014). The Value of Smarter Teachers: International Evidence on Teacher Cognitive Skills and Student Performance. NBER Working paper 20727.

Hanushek, E. A., & Rivkin, S. G. (2006). *Handbook of the Economics of Education*, Volume 2, Amsterdam: North Holland, 1052-1078.

Huitfeldt, I., Kirkebøen, L., & Rønning, M. (2016). *Effektevaluering av intensivopplæringen i Overgangsprosjektet, Ny Giv*: Andre delrapport. Retrieved from https://www.ssb.no/utdanning/artikler-og-publikasjoner/_attachment/258452?_ts=1532cdbfd00

Hægeland, T., Raaum, O., & Salvanes, K. G. (2005). Pupil achievement, school resources and family background. Retrieved from http://ftp.iza.org/dp1459.pdf

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

Jackson, K. C. (2010). Do Students Benefit from Attending Better Schools? Evidence from Rule-based Student Assignments in Trinidad and Tobago*. *Economic Journal*, 120 (549): 1399-1429. DOI: 10.1111/j.1468-0297.2010.02371.x.

Jackson, C. K., & Bruegmann, E. (2009). Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers. *American Economic Journal: Applied Economics*, 1:4, 85–108. DOI: 10.1257/app.1.4.85

Jepsen, C. (2005). Teacher characteristics and student achievement: evidence from teacher surveys. *Journal of Urban Economics* 57, 302-319. DOI:10.1016/j.jue.2004.11.001

June C. R. & Sanders, W. L. (2002). Teacher Quality and Equity in Educational Opportunity: Findings and Policy Implications. In *Teacher education*, edited by Williamson F. Evers, and Lance T. Izumi. Chicago: Hoover Institution Press.

Leuven, E., Oosterbeek, H., Rønning, M. (2008). Quasi-Experimental Estimates of the Effect of Class Size on Achievement in Norway. *IDEAS Working Paper Series from RePEc*.

Machin, S, & Salvanes, K. G. (2016). Valuing School Quality via a School Choice Reform. *The Scandinavian Journal of Economics*, 118(1), 3-24. DOI: 10.1111/sjoe.12133

Nechyba, T. J. (2006). Chapter 22 Income and Peer Quality Sorting in Public and Private Schools. In *Handbook of the Economics of Education*, edited by E. Hanushek, and F. Welch, 1327-1368. Elsevier.

OECD.org. (2016). "Education at a Glance 2016 - OECD Indicators. Retrieved from http://www.oecd.org/edu/education-at-a-glance-19991487.htm

regjeringen.no. 2016. *Slik blir den nye lærerutdanningen*.Retrieved from https://www.regjeringen.no/no/aktuelt/slik-blir-den-nye-larerutdanningen/id2503270/

regjeringen.no. (2016). Program for bedre gjennomføring i vidregående opplæring. Retrieved from https://www.regjeringen.no/no/tema/utdanning/grunnopplaring/innsiktsartikler/Bedre-gjennomforing-i-videregaende-/id2005356/.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica*, 73 (2): 417-458. DOI: 10.1111/j.1468-0262.2005.00584.x.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review,* 94 (2): 247-252. DOI: 10.1257/0002828041302244

St.meld. nr.44. (2008-2009). *Utdanningslinja*. Retrieved from https://www.regjeringen.no/no/dokumenter/stmeld-nr-44-2008-2009-/id565231/

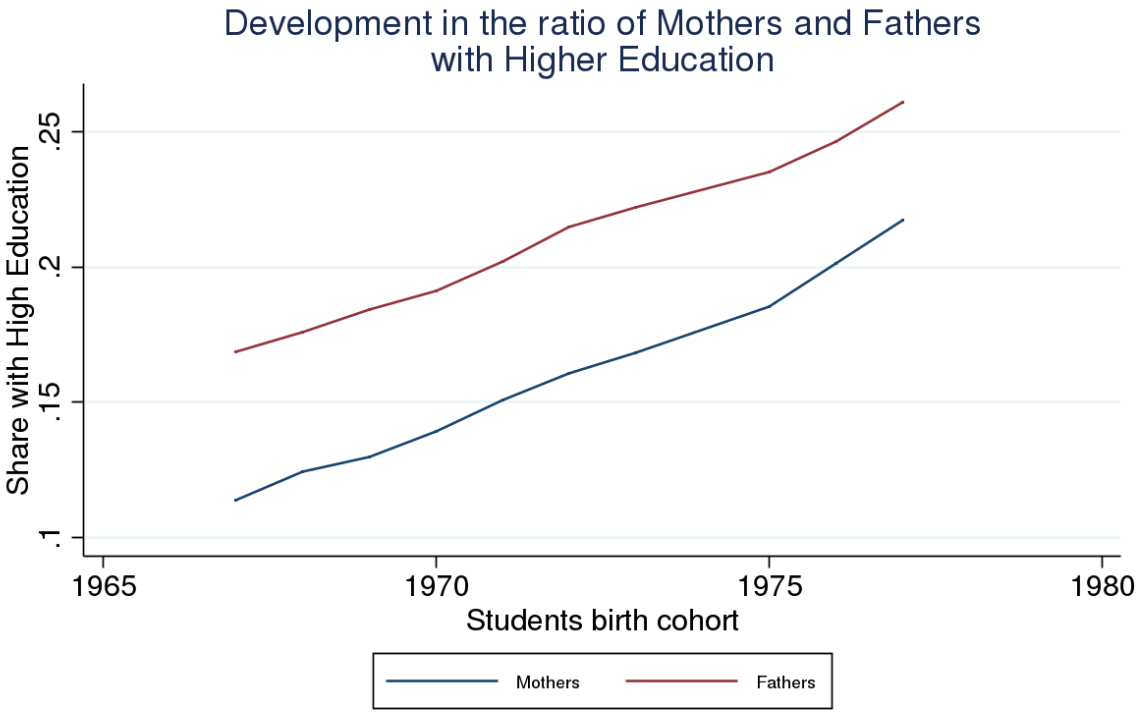
Statistics Norway. (2017). *Consistent disparities in national tests*.Retrieved from https://www.ssb.no/en/utdanning/statistikker/nasjprov/aar/2017-01-31

Stock, J. H., & Watson, M. W. (2015). *Introduction to Econometrics, Update, Global Edtion*. Updated third edition. ed. Pearson series in economics.: United Kingdom: Pearson Education M.U.A.

Wiswall, M. (2013). The dynamics of teacher quality. *The Journal of Public Economics,* 100: 61-78. DOI: 10.1016/j.jpubeco.2013.01.006

Appendix

Table A1. Correlation Matrix



Note: This graph plots the share of the parents with higher education by the bith cohort of their children, for all the ca 450 municipalities on wich we have data, exept for the birth cohort 1974 that lack many registrations

Table A2. Correlation Matrix

	High school on time	Share of Middle School Teachers with Long higher education	Share of Middle School Teachers with Teacher education	Fathers with higher education	Mothers with higher education	No. of minorities	Share full time teachers	Average teacher age	Teachers age squared	Class size
High school on time	1									
Share of Middle School Teachers with Long higher education	0.0261	1								
Share of Middle School Teachers with Teacher education	-0.0100	-0.5857	1							
Fathers with higher education	0.2092	0.0923	-0.0927	1						
Mothers with higher education	0.1661	0.0663	-0.0613	0.4235	1					
No. of minorities*	0.0184	0.0541	-0.1027	0.0413	0.0379	1				
Share full time teachers	-0.0254	-0.1105	0.0608	-0.0690	-0.0603	0.0065	1			
Average teacher age	0.0995	0.0832	0.0610	0.1148	0.1046	0.1199	-0.0975	1		
Teachers age squared	0.0970	0.0803	0.0574	0.1143	0.1036	0.1201	-0.1022	0.9956	1	
Class size**	-0.0173	0.1489	-0.1871	0.0321	0.0035	0.0817	0.0275	-0.0022	0.0020	1

Notes: Correlations on the sample with 452 middle schools in 256 municipalities, with 371,539 students, for the cohorts 1967 - 1977.Teacher characteristics are aggregated at middle school and cohort levels. *No. of minorities per cohort, in 7th grade. * *Total class size per cohort, in 7th grade.

Table A3. Regression Results

Dependent variable: Completion ratio at high school						
	OLS			**Municipality Fixed Effect**		
	(1)	(2)	(3)	(4)	(5)	(6)
Share of Middle School Teachers with Teacher education	0.0008	0.0008	0.006	0.0111	0.0111	-0.0008
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Fathers education		0.1968***	0.1926***		0.1926***	0.1912***
		(0.00)	(0.00)		(0.00)	(0.00)
Mothers education		0.1168***	0.1140***		0.1146***	0.1140***
		(0.00)	(0.00)		(0.00)	(0.00)
Minorities			-0.0022***			-0.0025***
			(0.00)			(0.00)
Full time			-0.0628***			-0.0644***
			(0.02)			(0.02)
Teachers age			0.0279***			0.0173***
			(0.01)			(0.01)
Teachers age ^{2}			-0.0003***			-0.0002***
			(0.00)			(0.00)
Class size			0.0001			0.0003***
			(0.00)			(0.00)
Adjusted R-Squared	0.0596	0.0596	0.0611	0.0669	0.0659	0.0665
N	371539	371539	371539	371539	371539	371539

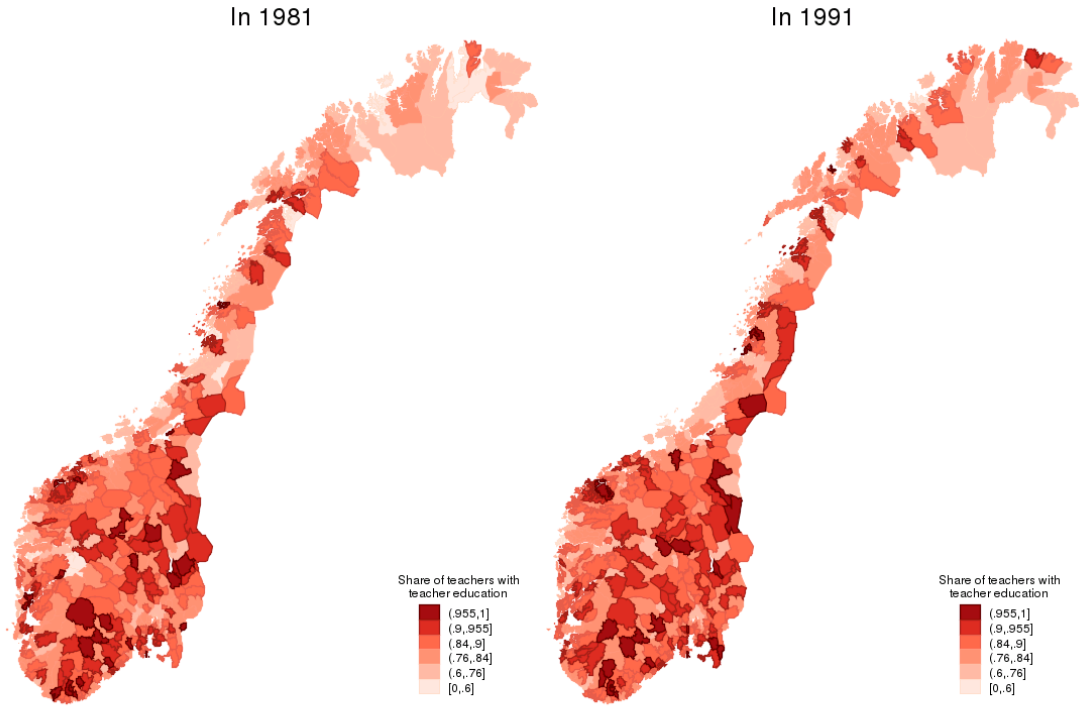
Notes:
Significance levels * p < .1; ** p< .05; *** p < .01. Estimated results of the OLS and municipality fixed effect models.
Middle school clustered standard errors in parenthesis. All regressions control for birth cohort of the students (dummies).
Mothers and fathers share with higher education are dummies on individual level of the students, while all other control variables are the share on school and cohort levels.

Table A4. Regression Results

Dependent variable: Completion ratio at high school						
	OLS			**Municipality Fixed Effect**		
	(7)	(8)	(9)	(10)	(11)	(12)
Share of Middle School Teachers with Long higher education	0.1191***	0.1185***	0.0703**	0.0960**	0.0958**	0.0811**
	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)
Fathers education		0.1954***	0.1920***		0.1924***	0.1911***
		(0.00)	(0.00)		(0.00)	(0.00)
Mothers education		0.1160***	0.1137***		0.1145***	0.1140***
		(0.00)	(0.00)		(0.00)	(0.00)
Minorities			-0.0023***			-0.0025***
			(0.00)			(0.00)
Full time			-0.0596***			-0.0640***
			(0.02)			(0.02)
Teachers age			0.0268***			0.0164***
			(0.01)			(0.01)
Teachers age ^{2}			-0.0003***			-0.0002**
			(0.00)			(0.00)
Class size			0.0001			0.0002***
			(0.00)			(0.00)
Adjusted R-Squared	0.0598	0.0598	0.0612	0.0660	0.0660	0.0665
N	371539	371539	371539	371539	371539	371539

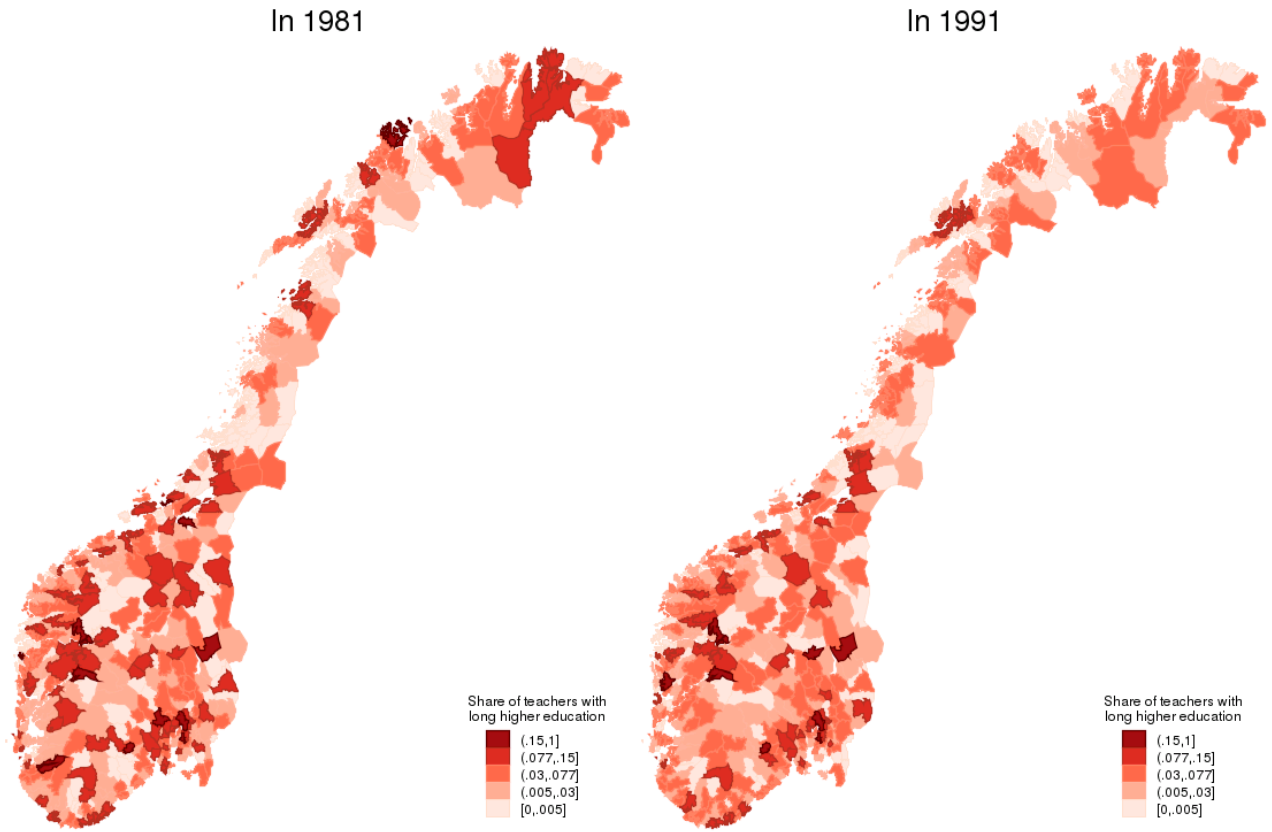
Notes:
Significance levels * p < .1; ** p< .05; *** p < .01. Estimated results of the OLS and municipality fixed effect models.
Middle school clustered standard errors in parenthesis. All regressions control for birth cohort of the students (dummies).
Mothers and fathers share with higher education are dummies on individual level of the students, while all other control variables are the share on school and cohort levels

A5. Distribution of Teacher Education



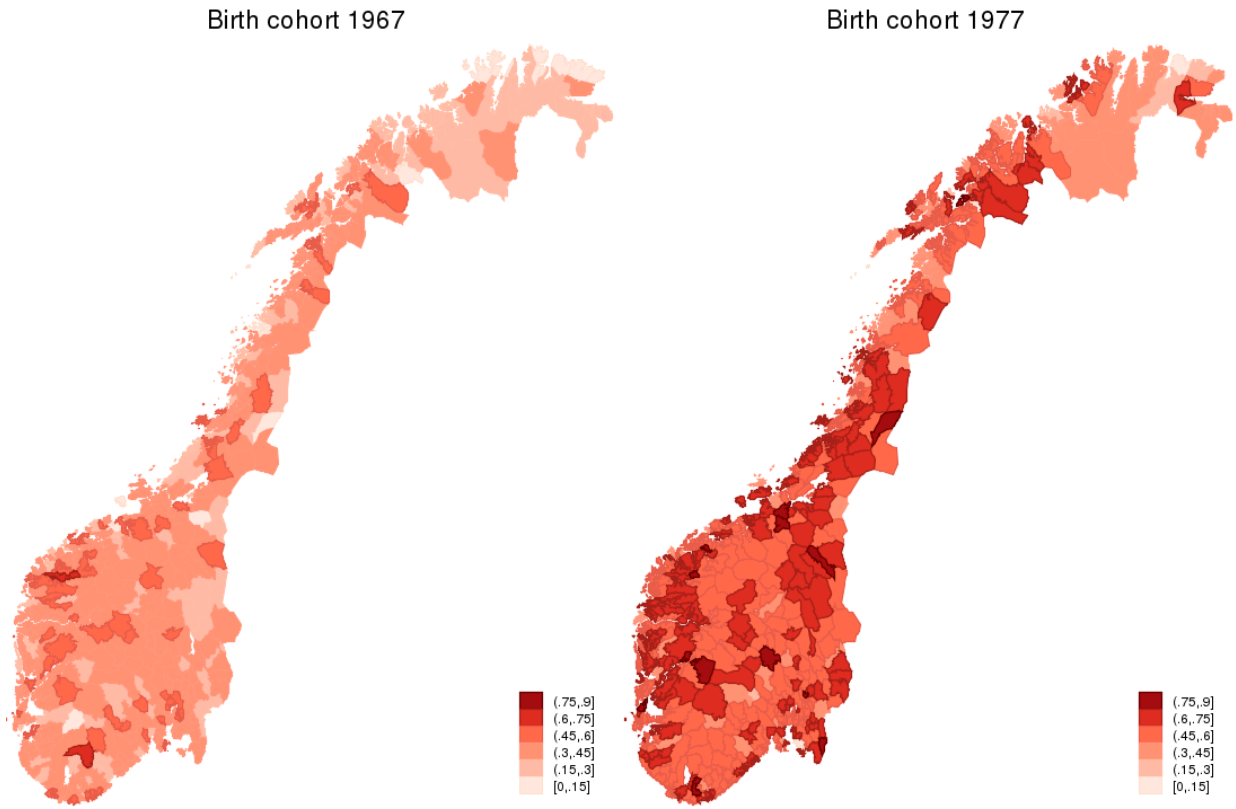
In these graphs we plot the teachers at middle school holding a teacher education in year 1981 and 1991 respectively,
for all the ca 450 municipalities on which we have data. The plot is based on the aproximate average percentiles. Municipalities colored in a darker shade have a larger share of teachers with teacher education.

A6. Distribution of Long higher education



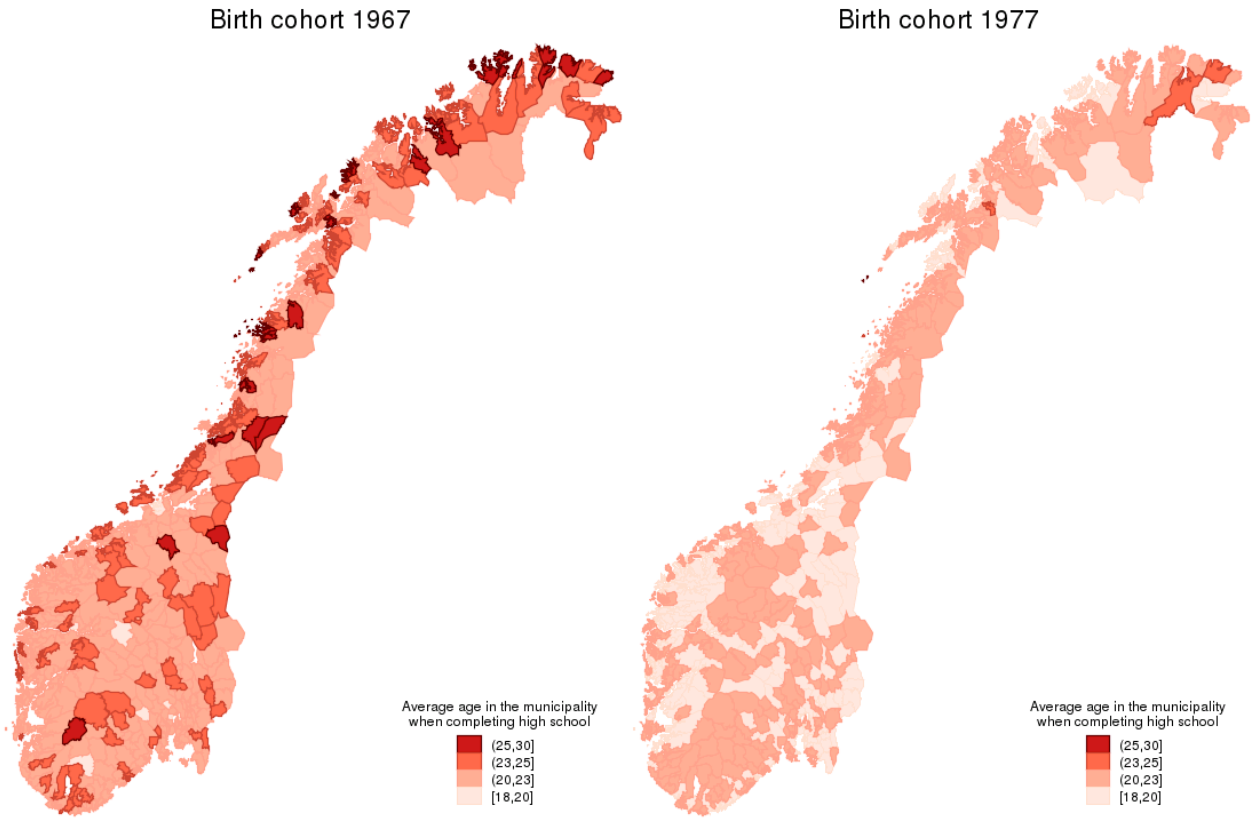
In these graphs we plot the teachers at middle school holding a long higher education in year 1981 and 1991 respectively, for all the ca 450 municipalities on which we have data. The plot is based on the aproximate average percentiles. Municipalities colored in a darker shade of red have a larger share of teachers with long higher education.

A7. Distribution of Students Completing High School on Time



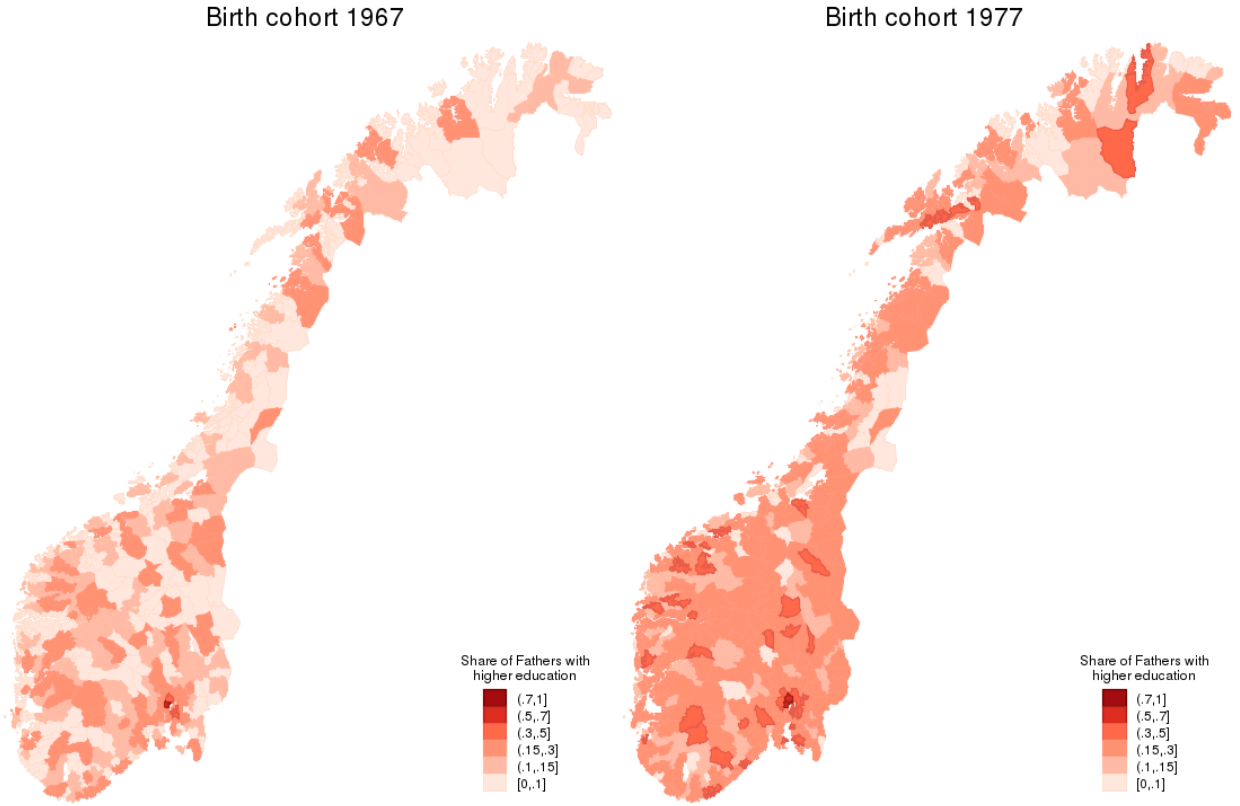
In these graphs we plot the high school completion ratio on time for the student birth cohort 1967 and 1977 respectively, for all the ca 450 municipalities on which we have data. Municipalities colored in a darker shade of red have a larger share of students who complete high school on time.

A8. Distribution of Students Age when Completing High School on Time



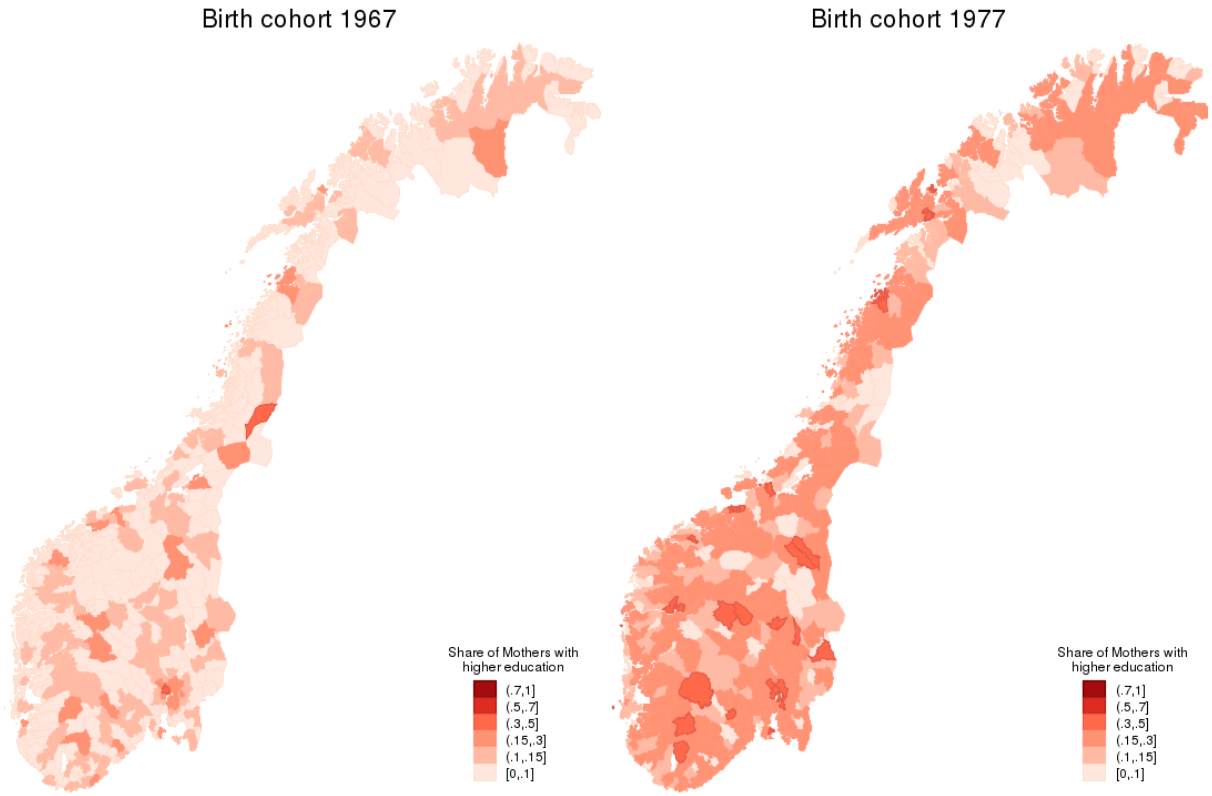
In these graphs we plot the age of the students when they complete high school for the student birth cohort 1967 and 1977 respectively, for all the ca 450 municipalities on which we have data. In municipalities colored in a darker shade of red, students have a higher average age when completing high school.

A9. Distribution of Fathers with Higher Education



In these graphs we plot the students fathers with higer education, for the student birth cohort 1967 and 1977 respectively, and for all the ca 450 municipalities on which we have data. Municipalities colored in a darker shade of red have a higher share of fathers with higher education.

A10. Distribution of Mothers with Higher Education



In these graphs we plot the students mothers with higer education, for the student birth cohort 1967 and 1977 respectively, and for all the ca 450 municipalities on which we have data. Municipalities colored in a darker shade of red have a higher share of mothers with higher education.