Returns to Field of Study: Evidence from a Norwegian Reform of College Expansion

Tora Kjærnes Knutsen

Master of Philosophy in Economics

Department of Economics

University of Oslo

Submitted May 10, 2017

Returns to Field of Study: Evidence from a Norwegian Reform of College Expansion

http://www.duo.uio.no

RETURNS TO FIELD OF STUDY: EVIDENCE FROM A NORWEGIAN REFORM OF COLLEGE EXPANSION

Abstract

Does the location of a college in it-self lead to higher education levels and increased earnings for people living nearby? I empirically address this question using an expansion of colleges in Norway approximately between 1960 and 1980. I argue that this expansion lead to exogenous variation in access to higher education. The new colleges offered specific fields of study: nursing, engineering, teaching and business. I draw on previous literature linking distance to college and degree attainment, but with additional variation across time and also field of study. I first establish a relationship between college establishment and degree attainment. Furthermore I use this to see whether these degrees gave increased wages for those with local access to a college, compared to those without. The results reveal that nursing degrees appear to have had a positive effect on female wages. For the other degrees the effect on earnings is not significant, but there are some indications that those whose parents had relatively high education experienced higher returns than other groups. Since the reform aimed at expanding the supply of education, the analysis may further throw light on the relative influence of demand side and supply side factors when it comes to educational choice. This thesis also opens up for more use of Norwegian expansion of higher education in a quasi-experimental framework.

Preface

I would like to thank Statistics Norway for providing me with Norwegian register data on education and income that has been essential for this thesis. I am also grateful to my supervisor Jo Thori Lind, who has shared great insight on empirical work and thesis writing. The work was done as a part of the project: "Inside the Black Box if Skill Formation: Determinants and Outcomes of Field of Study Choices" lead by Marte Rønning (Statistics Norway), who together with Jørgen Modalsli (Statistics Norway) have been of tremendous help as co-supervisors. The support and help from supervisors have made the work with this thesis both interesting and very stimulating. Their help has been crucial for this thesis. Crucial was also the help from those who took their time to proofread: Sigrid and my mother, Unni. However, all remaining errors remain my own. In addition, I would like to thank Statistics Norway for financial support and office space and ESOP for granting me their master scholarship. An essential reason for why I have enjoyed writing this thesis are (long) lunch breaks with other master students at Statistics Norway, I therefore want to thank them for their great company.

Tora Knutsen, May 2017.

Contents

List of Tables

List of Figures

Page intentionally left blank

1 Introduction

In Western countries higher education was subject to a very rapid expansion beginning in the 1960s, and developing through the 1970s and 1980s. This has later been called "the transition from elite to mass higher education"[1]. In Norway, this transition happened through an expansion of colleges, and not in the established universities. Education institutions were established among the country's 19 counties. Schools for engineering, nursing, teaching and a new type of college, regional colleges, were included in this major expansion, which for the most part happened outside the biggest cities. In this period policies that aimed at equalizing social and economic differences across the country, gained wide popular support. Thus it was politically stressed that the location of colleges was a measure for regional policy. Therefore the location of colleges was decided politically, and was not necessarily referring to local demands or resources. This provides us with a possible exogenous source of variation in access to college education both across time, space and field of study.

In Norway geographical background was important in predicting education, but the educational advantage by geographical centrality was steadily shrinking during the period after 1960 (Lindbekk, 1998). A part of the explanation could be the geographical distribution of colleges. The educational expansion happened in a period when Norway experienced a decrease in income inequality (Aaberge et al., 2016). The expansion of access to education at lower and higher levels has been pointed out as a possible explanation for this decrease (Pekkarinen et al., 2017). Therefore it is of particular interest to see which social groups that were affected by this reform. Were youth whose parents had high education, and who would have entered higher education anyway, induced to undertake a college education closer to home? Or were youth from less privileged homes, that would otherwise not have pursued college degrees, "pushed" into higher education? Another important change during this period was women's entry into the workforce. Therefore the gender aspect is important. How were women affected by the educational expansion and did it affect their labour force participation?

Using rich Norwegian register data on education I am able to track all individuals from the

[1]Identified by American sociologist Martin Throw in 1970.

areas where the new colleges were established and their educational outcomes in order to answer the question: Did the college expansion increase the chance of people living near a recently established college to undertake specific college degrees? A fixed effects model analysis[2] is used to answer this question. I also identify how the reforms affected different social groups, women and men, by using data on parental education and gender.

Considering the geographic expansion of higher education and the political process behind the location of colleges, I argue that the college establishments offer exogenous variation in degree attainment. In other words college establishment is assumed to be uncorrelated with other factors that might affect income, such as ability. Therefore I can apply an instrumental variable approach after linking degree attainment to college establishment.

Traditionally, years of schooling have been used to measure human capital and determine returns to education. However, in higher educations also field of study is important for labour market outcomes. Differences in returns across field of study rival the college wage premium (Altonji et. al, 2012) and therefore contributes to between-group inequality. An example is the persistent gender wage gap. Although women in many countries have surpassed men in higher education attendance, earning differentials between men and women are still pronounced. One explanation lies in different field of study choices across gender. At the college level, differences in field of study have received much less attention in the academic literature than the average return to higher education. This college reform, that affected specific fields of study, offers a unique possibility to develop a framework to study returns and choice of educational fields. Instrumenting college degree with local access to college at the age of 25, I explore whether the individuals that were "pushed" into the specific fields of study, got any monetary gains compared to others living in unaffected areas.

This thesis proceeds in the following way: Section 2 places this study in the context of relevant literature, section 3 gives a quick review of the Norwegian context, section 4 introduces the data and section 5 the empirical strategy. The following sections (6-7) present the first and second stage results. The final section (8) makes some concluding remarks returning to initial questions and some suggestion for further research using this college reform.

[2] All data analysis is done using Stata 13.1. When relevant I refer to which Stata programs that are used.

2 The Search for a Causal Effect of Education on Earnings

The estimation of the returns to education, both private and individual, has been the focus of considerable debate in the economics literature. A reoccurring issue is the relationship between education and earnings. There is a well-known positive relationship between years of schooling and earnings. The average return to a year of schooling from an Ordinary Least Squares (OLS) regression, is around 6 % internationally, although lower in Norway and the other Nordic countries (Harmon et al, 2000). However, there still remains some doubt concerning the causal effect of education on earnings. Does the education system provide input that increases an individual's earning potential or does the education system act merely as a sorting-mechanism? The screening hypothesis suggests that inter-educational earnings differentials reflect no direct productivity enhancing effects of education, but that education is a device that signals preexisting ability differences (Layard and Psacharopoulos, 1974) .

One of the most widely used equations to explain wage as a function of schooling and experience, is the one named after Jacob Mincer (1974). Typically, the logarithm of earnings is regressed on years of education and other control variables including years of labour market experience:

$$lnW = \alpha + \beta YearsOfEducation + \gamma Experience + Controls + u,$$

This procedure cannot measure the causal effect of education on wage because it ignores that one or more of the explanatory variables may be endogenously determined. Individuals differ in other characteristics that both determine education and earnings, such as ability. Unless we believe that we have controlled for all these characteristics, the coefficient of years of education captures economic returns also to these omitted variables, as well as the pure causal effect of education. Another issue is heterogeneous returns to education. If the returns to education varies across individuals the β should be interpreted as the average return to a year of education.

According to Becker's classical framework (1964) individuals reach an optimal schooling decision by balancing the benefits of schooling, which are realized over the life cycle, against the costs, which are born early on. Thus it is assumed that individuals seek to maximize the discounted present value of earnings and net schooling costs. However, individuals may have different abilities and tastes for schooling relative to work, and this variation may lead to differences in the optimal schooling level across individuals. As a consequence the economic benefits of schooling will differ and most likely create a positive upward bias in an OLS regression: People with a higher expected return to education will choose more schooling. On the other hand, people who would have made a lot of money anyway, have a higher alternative cost to schooling because they give up a higher income while going to school (Griliches, 1977). If this is the case, the coefficient in the OLS regression is biased downwards.

Card (2001) allows the optimal schooling to differ, not only because of ability, but also differences in discount rates across individuals. Individuals with a lower discount rate choose more schooling. One might suspect then that individuals with high-ability parents, who tend to offer children more resources for education, will have a lower discount rate, indicating that they are less impatient. A geographical expansion of education institutions may then affect those with a high discount factor the most because they react the most to a decrease in the cost of schooling.

Due to the substantial problems related to the Mincer equation; unobserved characteristics correlated with both wage and education as well as heterogeneous returns to education, alternative empirical approaches to measure the causal effect of education are widely used. A method to investigate this issue is to find exogenous variation that determines length of education, but that can be assumed to be unrelated with ability or other non-observed characteristics. Therefore randomness in assignment to education is commonly used as a natural experiment or as an instrumental variable.

The validity of an instrumental variable (IV) depends crucially on the assumption that the instrument is uncorrelated with other latent characteristics of individuals that may affect their earnings. Many studies have used educational reforms or distance to college as such exogenous variation. Taken as a whole the findings from the recent IV literature point to a causal effect of education that is as big or bigger than the OLS estimated return, at least for people whose schooling choices are affected by the instrument. One interpretation of this finding is that marginal returns to education among the low educated subgroups typically affected by supply-side innovations tend to be relatively high, reflecting their high marginal cost of schooling rather than low ability that limits their return to education (Card, 2001).

Natural experiments - Exogenous increases in length of education

Perhaps the most known study of the relationship between distance to college and years of schooling can be found in Card (1995). His results suggest that men who grew up with a nearby college have significantly higher education than other men. Assuming that ability is independent of college proximity, the distance to college can be used as an instrumental variable for educational attainment. He finds that education and earning gains are concentrated among men with poorly educated parents, men who would otherwise stop schooling at the lower levels. However it can be argued that living close to a college is not uncorrelated with ability, and therefore distance to college affects earnings through other channels than education. For example, families whose children are more likely to enroll in college may choose to locate close to a college (Heckman et al., 2006). The Norwegian college reform and Norwegian register data enable me to overcome this issue. This thesis uses partly the same instrument exogeneity assumption as Card, that distance to college is independent of ability, but with an additional source of exogeneity: we observe people in the same area before and after a college establishment.

Others have used institutional changes in the education system as exogenous determinants of schooling outcomes to help identifying the causal effect of education. An example is an increase in the years of compulsory schooling that is arguably unrelated to other characteristics, and can therefore be thought of as a “natural experiment”. This method is used both to measure the returns to education, as well as the effect of parental education on the outcomes of children (Black and Devereux, 2010). While studies of school reforms in for example Sweden and Finland show that reforms lead to increased incomes by increasing earnings of children with low-educated fathers (Pekkarinen et al., 2009 and Holmlund, 2008), a causal effect of parental education on offspring's schooling has been more difficult to establish. An example, studying educational mobility is Black et al. (2003)who use the increase of compulsory schooling in Norway from 7 to 9 years, which was implemented in different municipalities at different times, as an instrument for parental education. Using Norwegian register data, they find little evidence of a causal relationship between parental education and children's education, suggesting that the strong OLS correlation is due to family characteristics and not education spillovers. Other studies find some evidence that increases in parent's compulsory schooling decreases the probability that their child repeats a grade (Black and Devereux, 2010). These are studies of a reform at the bottom end of the education system and affect only those that would not otherwise have stayed in school.

Studies of reforms in higher education are less widespread. Maurin and McNally (2008) use the 1968 student revolts in France as a natural experiment. The period of protests coincided with a time in which important examinations took place. As a result, normal examination procedures were abandoned and the pass-rate increased in that year, therefore acting as an expansion of higher education. This affected in particular the cohorts 1948 and 1949 which are shown to have pursued further education because thresholds were lowered. Their results show a strong causal relationship between length of education and labour market outcomes. This effect is transmitted to the next generation, as children with parents from these cohorts are less likely to repeat a year in school. The affected are students from middle-class backgrounds since those with lower socioeconomic background were not likely to reach these entry exams to the university and students with high socioeconomic background had a higher pass rate in any given year. This study suggests that enabling 'marginal' persons to enter higher education can result in high private returns in the labour market and a long term effect, as this is transmitted to the next generation.

This thesis uses traits from these studies in the sense that geography and educational reform are used as exogenous variation in attainment of higher education. However, instead of focusing on length of education, I address some specific fields of study.

Field of study

Years of education has traditionally been used as a general measure of competence. This is a reasonable approach when there is little specialization in school, such as in elementary school. However, this is an incomplete measure when studying higher education, since the field of study also represents an investment in specific human capital. This kind of specific competence is one of the explanations that inequality has increased the last decades (Altonji et al., 2012). Field of study also contributes to inequality between groups such as the persistent gender wage gap.

As countries grow richer, gender disparities may shift from enrollment and learning to segregation in fields of study (World Bank, 2012). Thus, to understand the persistence of this wage gap, the focus of study has to change accordingly. The Norwegian labour market is gendered in dimensions such as sector, industry and profession. In Norway, women earn 85 %[3]of what men earn. This gap has been stable over the last decade. The wage disparities are largest among people with higher education, especially for people with short higher education from universities or university colleges[4]. In the latter category, women earn 79 % of what men earn. One explanation is that women and men chose different fields of education particularly at this level in the education system. Women often choose health related fields and work in the public sector while men choose technical or scientific fields and work in the private sector (Statistics Norway, 2009). This provides some motivation for why expansion of colleges can be used to study the gender wage gap.

One of few studies looking at field of study is Kirkebøen and co-authors (2016), who use information from the process of centralized admission to higher education in Norway. Applicants rank their choices and are admitted based on their grade point average. This creates instruments from discontinuities that randomize applicants near unpredictable admission cutoffs, picking people that are on the margins of entry to a particular field. This identification strategy allows for a measure of the payoffs of completing one type of education relative to a

[3]In this comparison the wage of people working less than full time is converted to the wage of a full time employee. Thus, this is not considering the wage differences due to differences in working hours between men and women.

[4]This pattern remained unchanged between 1998 to 2008 (Statistics Norway, 2009)

particular next-best alternative, and by this identifying the differences in payoffs between the two. They find that field of study drives the heterogeneity in the payoffs to post-secondary education and that the gender pay gap is relatively small compared to earning differences by educational field. The effect of earnings from attending a more selective institution tends to be relatively small compared to the payoffs of field of study, although this might be particular to the Norwegian context. They find a small positive payoff to attending a college compared to a university, when keeping field of study fixed. The authors conclude that individuals choose fields in which they have a comparative advantage.

Their data on applicants ranking shows that individuals rank related fields of study. A majority of the applicants who prefer engineering rank science and technology as their second field. Those who prefer health, rank social science and teaching as their second best options. Those ranking business as their preferred choice seem to be a more heterogeneous group, ranking both social science and natural sciences as second bests. This demonstrates a preference pattern that might be interesting for this thesis: The establishment of a new field of education is only important for individuals who have an interest in the field offered. If you prefer a technical field, you are not directly affected if a nursing education is established near you.

Field of study and the Norwegian college reform

The foregoing theoretical discussion suggests an ambiguous effect of college establishment on degree attainment for people living near a college. On the one hand, decreasing the distance to college should affect those who would otherwise have stopped schooling at lower levels. Because a decrease in the cost of schooling, in the sense that the travel cost diminishes, should affect people with low discount factors, and by that little "patience" for education. On the other hand, individuals with higher returns to schooling choose more schooling. These individuals choose schooling anyway and may choose a field of study closer to home, but not more schooling than in the counterfactual scenario without a college establishment. If these are the only people entering colleges, there may be no detectable effect. When it comes to earnings, both theory and empirical evidence mostly suggest that the effect of education on earnings should be positive. Though if people entering the new colleges would have entered college anyway, there may be no effect on earnings.

3 Background

Expansion of higher education in Norway

In the years after WW2 the priority was rebuilding the country, not expanding education. The priorities changed in the nineteen-fifties when education at primary school level was reformed. The reforms culiminated in an increase in mandatory schooling from 7 to 9 years, a transition implemented gradually between 1960 to 1975. The primary effect of this reform was to reduce the number of people with less than 9 years of education. In addition a unified high school system was introduced with more generous admission standards than before (Black et. al, 2003).

The next phase of reforms was in higher education. From the 1960s to the 1980s the structure of higher education in Norway changed dramatically. While the number of students at the universities remained stable around 40 000, the college sector grew from 67,000 to 90,000 students in the ten years between 1975 and 1985 (Johnsen, 1999). The percentage of the population older than sixteen with higher education at college or university level more than doubled between 1970 to 1990, increasing from 7% to 16%. Figure 1 [5]shows that the largest increase between 1970 and 1990 happened in short tertiary education, which was the type of education offered in the college sector.

[5]Table from Statistics Norway available from https://www.ssb.no/en/statistikkbanken.

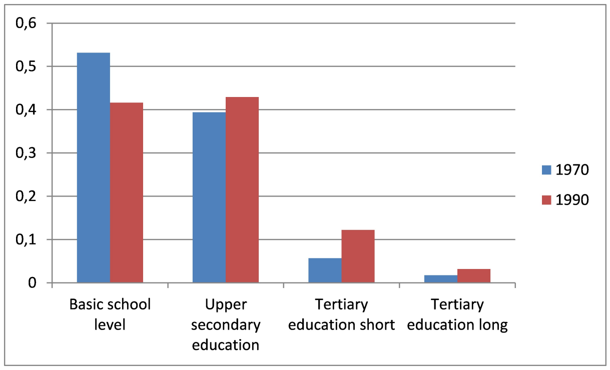


Figure 1: Education level of the population in 1970 and 1990
(Table from Statbank, Statistics Norway)

The geographical distribution of colleges

In the nineteen-sixties, as the expansion of higher education intensified, Norway was divided into thirteen regions based on the counties of Norway [“fylke”][6]. The new educational institutions were mostly distributed among these education regions. Regional college boards [regionale høgskolestyrer] were established in order to integrate teaching, engineering, nursing and other educations at college level into a regional system of higher education. Politically it was stressed that the location of higher educational institutions was a regional policy measure (Norwegian Ministry of Education, 1975).

During the seventies, some regions were divided and the counties became the education regions. A new type of college, regional colleges, were established in 12 locations across the country. These colleges offered varying degrees depending on location, and most offered a degree in business. After the expansion most counties had at least a nursing college, a teaching college, an engineering college and a regional college or a university. The complete overview of the establishment of these education institutions is given in Appendix A. The final integration of all college education into regional units did not happen until the next important

[6]These regions were: Troms/Finnmark, Nordland, Trøndelag, Møre og Romsdal, Sogn og Fjordane, Hordaland, Rogaland, Agder, Telemark, Oppland/Hedmark, Oslo/Akershus, Vestfold and Østfold

reform in the Norwegian higher education system. In 1994, 98 colleges were merged to 26 larger units (Hernes, 1991).

The Norwegian education system

Education in Norway consists of mandatory elementary school from first grade to grade 9[7] and 3 years of high school[8]. In high school, students can choose between various vocational tracks and academic tracks. Only the academic tracks give access to college or university education. Further down I will explain that some types of education have changed levels in the education system over time. There have also been various possibilities of entering colleges without high school degrees in specific fields. In Norway, University College [høgskole], which I for simplicity call college, is the name of a level within higher education that traditionally has offered shorter more career-oriented types of education at what today is know as "bachelor level". Another characteristic difference between colleges and universities has also been that colleges do not conduct academic research.

A tendency particular for Norway in the decades after WW2 was that youth left school after elementary school (7-9 years) and came back into the education system later, in particular to colleges and vocational schools (Bostad, 2007). OECD identifies this as a feature that makes the Norwegian tertiary system distinctive: students are somewhat older when they commence and graduate than in many other countries (Clark and Sohlman, 2009). We will see that this is relevant for the analysis of the college reform as we cannot identify unique cohorts affected by the reforms.

Another distinctive feature of tertiary education in Norway is that it mainly relies upon public funding: Public universities and colleges have very low or no tuition fees. Thus the cost of education is that of covering living expenses and the alternative cost of not working. Funding is available to everyone, especially after loans and scholarships ceased to be means-tested on parental income in 1968/1969. After this a scholarship was given to all students not living

[7] Grade 10 after 1997.

[8] Before "Gymnas" and "Yrkesskole", in the 1970's transformed to "Videregående skole" .

with their parents. In 1969 only 3 % of the applicants did not get granted a scholarship and loan (Ministry of Education, 1969). Thus, during the period we are looking at, education was made more available both because financial support increased and was made more accessible, and that higher education institutions were distributed across the country.

Social background and educational expansion

The income mobility increased considerably for cohorts born between 1930 and 1950, and stabilized for cohorts born after 1950. The reason for this rather dramatic change is not fully understood but the development of the welfare state possibly played a role. Education policy could be an essential factor, with the expansion of elementary school and later the expansion in higher education according to Salvenes (2017) . However, according to the same author paradoxically parental education has almost as large impact on children's education in Norway as in the U.S. Thus, despite low income inequality and high social mobility in terms of income, educational inequality has persisted. Salvenes suggests two possible reasons for this: First, cultural norms and expectations are more persisting and are slowly changed by policy. Second, as the returns to higher education is lower in Norway compared to other developed countries, there are less incentives to undertake higher education.

An interesting question is whether the expansion of higher education leads to changes in the educational persistence between parents and their children over time. Banden and Machin (2004) find that between 1970-1990 in the U.K, the increase in participation in higher education among people from different economic backgrounds has not been equally distributed: Children from richer backgrounds (and possibly higher educated) benefited disproportionately from the expansion in higher education and the participation gap in education between rich and poor increased. Bratti and co-authors (2008) study an expansion in higher education in Italy in the 1990's and suggest that the expansion had only limited effects in terms of reducing educational inequality. Heineck and Riphahn (2009) find no significant change in the inter generational persistence in education in Germany the last five decades.

There is a rich literature in sociology on the development of social background's impact on higher education in Norway. Lindbekk (1998) finds that the reform in mandatory education raised general attainment levels. An equalization trend is therefore observed for cohorts born before 1965. The importance of class decreased, but the impact of parent's educational background on educational attainment remained stable. Aamodt (1982) examines the social selection taking place in the educational system and finds that the selection happening before high school was of much greater magnitude than the one happening after. He finds that among pupils with similar school achievements, but different social backgrounds, those belonging to higher social strata more often chose the secondary education track that lead to higher education.

Professions subject to this study

In order to study the effects of the reforms, I choose the professions that were the most affected by college expansion in this period (ca. 1960 - 1980). These are nursing, engineering, teaching and business and management[9]. The latter is in this thesis for simplicity called business. Nursing and engineering represent two widespread educations that are respectively female and male dominated. Business and teaching represent degrees with a greater gender mix. In the former, one third were women while in the latter, two thirds. The regional colleges were perhaps the educational institutions with the clearest objective of equalizing education across the Norwegian regions. Business was offered at most regional colleges and is therefore useful for the purpose of studying the effect of regional colleges. Figure 2 demonstrates the expansion of these degrees, showing the share of cohorts with the respective field of college education. The increase is quite dramatic for engineering and nursing[10].

[9] "Økonomi og administrasjon"

[10] Other historical sources tell similar stories. The documentation of censuses 1960, 1970 and 1980 (Vassenden, 1987) contains statistics on education. In the 1960 census education from type of institution was registered. In the 1970 and 1980- census, type of education was registered. As percentage of the population, the share of nurses increased from 0,59 % in 1960 to 0,98 % in 1970 and 1,3 % in 1980 . The share of college engineers increased from 0,5% in 1960 to approximately 1 % in 1980. Hence, the number of nurses and college engineers doubled as a percentage of the population between 1960 and 1980.

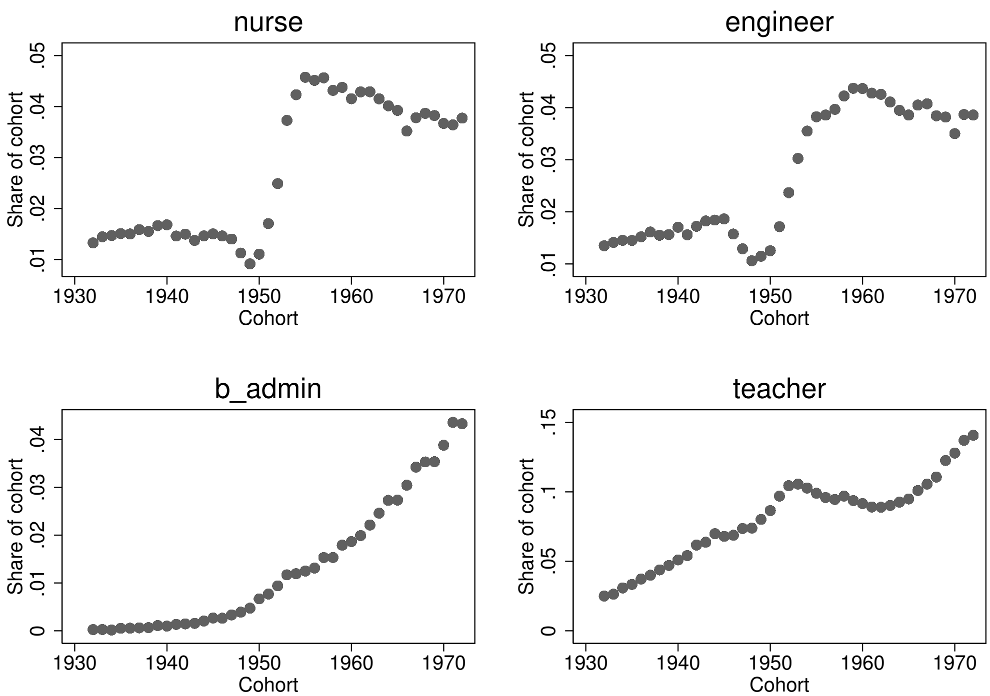


Figure 2: Share of cohort with nursing, engineering, business or teaching degree

Nursing

Nurses in Norway organized early; in 1912 the Norwegian Nurses Association was established. From the beginning they campaigned for a better and standardized training of nurses. As a result the education of nurses was a 3-year degree starting some places already in the 1920's and national standardized exams were introduced in the 1950's (Norwegian Nurse Association, 2017). Nursing was at the level of upper secondary schools until it was upgraded to university college in 1981. Nursing was and continues to be a profession dominated by women. In the data, as much as 88 % of nursing degree holders are women. Average age for initiating a nursing degree in the data is around 22 years[11]. Many cities have a history of

[11]We only have the year of education for those who started their degree after 1974.When upgraded to a university college, the average age changes slightly, from 22 years between 1974- 1980 to 23,6 between

educating nurses dating back to before WW2, so we may suspect that these cities had some particular characteristics. Therefore I look at new nursing colleges established after 1950 when education was already subject to government planning and geographical distribution was a main concern. In this period, new nursing colleges were established in 11 locations[12].

Engineering

In 1962, the education in technical schools increased from 1 to 3 years[13]. In 1977, the schools changed name from technical schools to engineering colleges and the criteria for entry changed; now a high school diploma was a prerequisite. Approaching 1990 most counties had an engineering college[14]. With a male percentage of 88, engineering is (surprisingly) exactly as dominated by men as nursing is by women. As mentioned previously, the college system was subject to big changes in the nineties. But for the engineering colleges, already between 1986-1990 the degree structure changed gradually and 3-year degrees for all were introduced.

Teaching

After WW2 there was a large shortage of teachers. The expansion of teacher education was coordinated by "The Teacher School Committee[lærerskolekomiteen]". In the 1950's it was not longer possible to further expand the already existing institutions. This led the government to establish new teacher colleges in the big cities[15]. Their location was highly dependent on local initiatives (Johnsen, 1999). In the 1960's it was assumed that there was a

---

1981-1990. There is no significant difference in average starting age between treated and untreated regions.

[12]Nursing colleges were established in the following cities (year of establishment in parenthesis): Fredrikstad(1955), Molde(1958), Hammerfest(1960), Gjøvik(1971), Ålesund(1974), Kristiansand(1976), Førde(1979), Stord(1977), Haugesund(1980) and Harstad(1982).

[13]2 years if already completed the science track[real/naturfagartium] in high school

[14]After World War Two new engineer colleges were established in the following cities: Hønefoss(1955) Narvik(1955), Sarpsborg/Tune(1965), Ålesund(1965), Grimstad(1967), Gjøvik(1966), Kongsberg(1970), Tromsø (1981, Maritim engineer), Haugesund(1984) and Førde(1988).

[15]New schools were established in Oslo(1945), Bodø(1951), Bergen(1953), Stavanger(1954), Tønsberg(1958) and Trondheim(1959). Oslo, Trondheim and Bergen had existing teacher colleges and are therefore not considered to be affected by the reform.

need for an increase of 25 000 teachers. The Parliament then decided to increase the number of students in existing schools (Ministry of Education, 1964). When this was not enough, new colleges were established in: Halden(1963), Sogndal(1973) and Alta(1973), which by then were located in the only three counties without a teaching college. The teaching was upgraded to a 3-year degree at college level in 1973.

Regional colleges

In 1969, a government committee (the so-called "Ottoesen committee") suggested the creation of regional colleges in a white paper to the parliament (Ottoesen, 1969). The regional colleges were to be distributed across the country. These colleges aimed to shift demand from longer university education to shorter and more carreer-oriented education. The parliament aimed at establishing one in each educational region that did not have a university, but this was for political and budgetary reasons not realized for all regions. The location of the regional colleges was mostly decided by the parliament, where this was subject to debate with conflict between the different regions.The idea of equalizing across regions appears to be important in the decision-making.

Most regional colleges[16] offered a 2-year degree in business. This was also the most popular education at the colleges (Johnsen, 1999). Since this degree was offered at most regional colleges, while other subjects differed due to local factors, this degree is the most promising candidate for exogenous variation. Also, the degree followed a standardized curriculum proposed to the parliament by the ministry of education[17].

Salvanes and co-authors (2014) use the expansion of regional colleges to measure the impact of an exogenous increase in the supply of skilled labour. Municipalities with regional colleges are compared to some "counter-factual" synthetic optimal control municipalities that share the same trend as the "treated" before the reform. They find that both the share of skilled

[16] With an exception of Lillehammer and Bodø.

[17] In White paper (St. prp.) nr 136 (1968-69), "Om prøvedrift med distriktshøgskolar [About the trial phase with regional colleges]"

labour and average skilled wages increase as a response and suggest that this is caused by a skill biased shift in demand.

4 The Data

Field of study

The data used to assess the effect of the college establishments on degree attainment is mainly Norwegian Education Data Base (NUDB). This data base contains register data on education for the whole population from 1970 and onward including all educations initiated (NUDB, 2007) and the parent's education for about eighty percent of the population. Education is coded with six digits so that we are able to identify particular educational fields. Fields that have changed levels, such as nursing, are coded with the level they have today.

A problem is that NUDB lacks education data for around 800 000 individuals, the large majority born before 1958. Around half of them are registered in NUDB, but without information on education. This is resolved using the censuses from 1960 and 1970 which also contains information on level and type of education. For these individuals highest achieved education in 1970 is used and the fields subject to this study are re-coded. Unfortunately this makes it impossible to detect if these individuals took the specified degrees and started education at a higher level later. Thus, we lose individuals who took a nursing degree and entered a higher level degree at a later stage. However, there is no reason to believe that many did. Another difference is that the NUDB records people that initiate a degree, but never finish, while the census only register the current education level.[18]Appendix B.1 shows the codes used to identify the different professions.

I have not been able to find information on how the fact that engineering colleges changed degree structure between 1986 and 1990, affected the coding. Therefore I only consider

[18]At an earlier stage working with the thesis I did the analysis using only individuals in NUDB and found qualitatively the same results.

engineering degrees taken before 1990. NUDB has every education initiated linked to a national identity number. This implies that we can track every individual's educational career. As a result we have in most cases more than one observation per individual. In order to obtain a data set with one observation per individual, the data is collapsed on the highest attempted education level and the respective field/degree we are looking at. If an individual has pursued other degrees in higher education, for example continued on a master program at a later stage, they are also included (for those whose information is gathered in from NUDB). Parent's education and municipality at the age of 16 are uniquely identified.

Table 1 displays the share of the whole population and the population by subgroups, with college degrees in the fields addressed in this thesis. Important to note is the group with no information on parental education level, who appears more similar to the group whose parents have only elementary school, than to the other groups. Other than for the business degree, the differences in means between the whole population including or excluding the big cities are not substantial. When split into subgroups according to the four categories of parental education level, we see that all degrees are more common among people whose parents have short higher education than among the other groups and that there are big differences in attainment between people whose parents have completed high school and those whose parents have not.

Income

To measure income, total pre-tax income from work is used.[19] These registers are available from 1967 and onward. To meaningfully assess earning differences we must keep work-experience roughly fixed. Therefore earnings have to be observed for the same age-group. I have chosen to use the income of individuals when they are 35-40 years because this should be a period when people are finished with their education. Therefore we can go back as far as to cohorts born in 1932, the cohort aged 35 years in 1967. Data on income between

[19]More specifically income that gives entitlement to pension ["pensjonsgivende inntekt"]. This includes mostly income from work, but also benefits that enter as a substitute to income from work such as paid sick leave and temporary disability benefits

Table 1: Summary statistics : Nursing, engineering, business and teaching degrees

	(i) **Nurse**		(ii) **Engineer**		(iii) **Business**		(iv) **Teacher**	
Population	*Mean*	*Sd*	*Mean*	*Sd*	*Mean*	*Sd*	*Mean*	*Sd*
All (N=2309176)	0,03	0,17	0,036	0,19	0,015	0,12	0,085	*0,028*
Women (N=1124659)	0,05	0,22	0,005	0,07	0,014	0,12	0,12	0,32
Men (N=1184517)	0,006	0,08	0,065	0,25	0,016	0,13	0,06	0,23
Excl. Oslo* (N=1746516)	0,03	0,17	0,035	0,18	0,017	0,13	0,09	0,029
Excl. Urban** (N=1275700)	0,03	0,17	0,035	0,18	0,017	0,13	0,092	0,29
By Parental ed.								
Missing *** (N=456774)	0,019	0,14	0,046	0,21	0,0016	0,04	0,052	0,22
Higher education								
4 years + (N=85017)	0,048	0,21	0,042	0,20	0,028	0,17	0,17	0,38
1-4 years (N=185192)	0,058	0,23	0,046	0,021	0,037	0,19	0,18	0,38
No Higher Education								
High School (N=900960)	0,036	0,19	0,035	0,18	0,022	0,15	0,10	0,31
Elementary School (N=684813)	0,017	0,017	0,027	0,16	0,0076	0.09	0,048	0,21

Note: The included populations is the Norwegian population born between 1932 and 1972.
*Excluding Oslo Region, i.e area around Oslo, defined by Norwegian Labour Market Regions (Bhuller, 2009).
**Excluding area around 4 biggest cities, Oslo, Bergen, Trondheim and Stavanger.
*** No registered information on parental education.

1967 and 2013 is merged with Norwegian register data on education (NUDB) and data from censuses. 5700 observations are not matched and are therefore deleted. Cohorts born between 1932-1973 consist of around 2,3 million individuals. This number is confirmed by adding the number of births between 1932 and 1973[20]. Around 160 000 individuals have missing or zero income at the age of 35 to 40 years, two thirds of these are women. Before 1977 the value missing is more frequent, a possibility is that this means zero income. The 160 0000 are dropped in the regression of college attainment on earnings. However I want to explore the non-earners because there is a possibility that for example a nursing college could increase labour supply also on the extensive margin. Women who would otherwise have been outside the labour market could get a college degree and start working. Therefore a separate regression on female labour force participation is done instead of including zero-earners.

Location and timing of establishment of colleges

The information on the location, timing and development of the expansion of colleges in Norway builds on documents from the Parliament: Both white papers and discussions in the parliament which are all digitized and available from the Parliament's website[21]. An in-depth analysis of the political process leading to the establishments of regional colleges is found in Johnsen (2009). The year of college establishments is gathered from Norwegian Center for Research Data (NSD) and Norwegian Research Council's publication on higher education institutions from 1979 (NAVF, 1979)[22].

Geographical divisions

The map of Norwegian municipalities changed substantially in the 1960s. Therefore the municipality borders of 1980 are used. The municipality of residence at age 16 is used

[20] Population, births and deaths from 1734 until today can be found here: http://www.ssb.no/a/histstat/tabeller/3-13.html

[21] www.stortinget.no

[22] Then called "Norges Almennvitenskapelige forskningsråd", now a part of The Norwegian Research Council.

where this information was available. This variable was recorded for the cohorts born in 1958 and onward. For the rest I have used the municipality of residence from the census in 1960. Students registered as living outside Norway (could be exchange students etc) and with unknown municipality are dropped. The latter constitutes of about 3 % of the population and 2/3 of this group are immigrants (people born outside Norway with foreign parents). Since this analysis is based on location it is not possible to include individuals without information on municipality.

In order to identify the effect of having a college nearby we need to define what is nearby. This must be a measure of where individuals could access without changing their place of residence. The municipalities would be too restrictive in many cases and the counties too big. What is a reasonable travel distances is also not obvious. The Classification of Economic Regions provides us with this exact information. Norway is divided into 90 regions based on data on the number of people commuting between different municipalities and some other more qualitative measures. A map showing Norway divided into Economic regions is shown in Appendix B.2. One limitation of this measure is that in order to correspond to Eurostat's regional level, the regions cannot cross the borders of the different administrative regions in Norway, the counties. A big city like Oslo is a region on its own since it is also county, even though there are certainly many people commuting to Oslo residing in the municipalities close to Oslo. However, since the reform in Norway's case was seen as an instrument to equalize educational differences across the Norwegian regions we are particularly interested in its effects outside the biggest cities.

To define the Oslo region I have used Norwegian Labour Market Regions[23]. Here Norway is divided into 46 regions and these are allowed to cross the county boundaries and are restricted to a minimum of inhabitants (17,500 in year 2000) (Bhuller, 2009).With this definition many regions in scarcely populated areas become large in geographic size. I use the 46- regions in regressions to see how the use of regional divisions matters. The classification of economic regions (90 regions) is based on data from year 2000. As infrastructure was less developed before than it is today, we might be exaggerating the possibility that people had to commute

[23]According to this definition the Oslo region includes all parts of Akershus, large parts of Østfold and some parts of Vestfold, Hedmark, Oppland and Buskerud.

within these regions. Thus the local effects would be underestimated if this is the case. A possible extension for further research is to use distance to college instead of a regional division in order to not make any assumption about travel time and accessibility.

5 Empirical Strategy

The empirical analysis follows a standard instrumental variable approach. Perhaps different than other usages of instrumental variables, is that this thesis is as much concerned with the question raised in the first as in the second stage. Although different from the classic empirical methods considered when reviewing literature in section 2, the model used can still be explained in those terms: We believe that there is a connection between wage and undertaking a college degree, $lnW = \beta CollegeDegree + Controls + u$, but we suspect that college degree is correlated with the error term because it is determined by unobserved factors that also determine wage. To overcome this problem of endogeneity, I use randomness in access to college degrees caused by the Norwegian college expansion.

The first of two key requirements for a valid instrument is instrument relevance. The instrument must be correlated with the endogenously determined variable, here college degree. Therefore, first stage regressions are performed to verify the first step: $CollegeDegree = \gamma AccessToCollege + Controls + \varepsilon$. Did the reform affect people living in regions where colleges were established? This is an important question in it-self because a goal of the reform was to increase the access to education for people living far from the already established education institutions. Did the reform result in access to higher education for the "masses"?

The second step is to use the fitted value $\widehat{Degree}$ from the first stage regression in the wage regression above. This enables me to identify a causal effect of college degree on earnings if the second requirement for a instrumental variable holds: The instrument must only affect wages through the college degree channel. This is not possible to explicitly verify, therefore I will discuss to what extent this assumption holds.

The indicator of social background used in this analysis is parental education level. Where the highest level (1) obtained by parents corresponds to one or both parents having more than four years of university education, and the lowest level (4) to primary school as their highest education[24]. An interesting question is which of these social groups were affected by a new college establishment. To explicitly look at how the reform affected individuals with different social backgrounds, I do the regressions on each subgroup. This will also give an indication of which group that drives the results.

Gender and parental education are pre-determined characteristics that affect degree attainment, but not college establishment. They can therefore be included as control variables, but are not necessary for causal inference. I do one regression controlling for both to make sure the effect is robust to background and gender. However, in the other regressions these controls are not included in order to make the interpretations as easy as possible. How is the interpretation, for example, of the the coefficient on nursing college establishment controlled for being a woman, when almost all that entered nursing were women?

5.1 First stage: Establishing a link between college establishment and college attainment

The first regression model connects attainment of college degree to access to college in home region in the following way:

$$\widehat{Degree}_i = \alpha + \pi AccessToCollege_i + \sum_{1932}^{1972} \lambda Cohort_i + \sum \gamma Region_g + \theta Controls_i + \varepsilon_{ig} \quad (1)$$

Where *College Degree* is an indicator variable equal to 1 if an individual has a degree in the specified field. *Cohort* designates a full set of cohort dummies while *Region* refers to region fixed effects. Controls are a set of social background dummies and a dummy for women.

[24]Primary school is defined as having finished education at level 1 or 2, which refer to “Barneskole” and “Ungdomskole” today, and “folkeskole” and “Framhaldsskole” before reforms in the sixties, this means 1-9 years of schooling. Level 3,-one or borth have high school education including the old “realskole” , refers to 1- 4 years of high school, thus more than 9 -13 years of schooling. Level 2 refers 1-4 years of higher education at college level.

Access to College is a dummy variable equal to 1 if an individual had access to a college in the specified field in his or her home region at the possible age of education. I will look at age of education being 19 and 25.The establishment of a new education institution can have drawn people of different ages to education. Table 2shows the results when using 25 years, whereas the results for 19 years is shown in Appendix C.3. The choice of 19 and 25 refers to respectively 19 being the age most finish high school, accounting for people entering a college at an early age, and 25 accounts for people entering later. 25 years appears to give the strongest results.

I do the regressions on the following levels: i) The baseline regression using the whole sample, ii) to see if the effect is robust to another geographical division I use the alternative regional division previously introduced (46 regions), and to examine the effects for subgroups I do the baseline regression iii) excluding Oslo area iv) separately by gender and separately by v) parental education level. In addition a separate regression controlling for gender and parental education is performed.

The error term, u_{ig}, is clustered at the regional level as the error term for individuals in the same region may be correlated. The inclusion of cluster-specific fixed effects generally does not control for within-cluster correlation of the error term. Using between estimation[25] gives the correct cluster robust standard errors as a regression with dummies would use the wrong degrees-of-freedom correction. One might worry that there is clustering within cohorts. However, if the within-cohort clustering is due to shocks that are the same across observations in a given year, then including cohort fixed effects as regressors will absorb this clustering (Cameron and Miller, 2015) .

This is a linear probability model (LPM). A concern with a LPM model is that it is not constrained by the interval between 0 and 1.This is not so problematic for the interpretation of the results as I am not working with any continuous variable and I therefore need not worry what happens at very large values of any variable. When it comes to marginal effects, the choice between non-linear models such as logit and probit and the LPM matters little. Also extra complexity comes in if nonlinear models are used to work with instrumental variables.

[25] Stata's xtreg- command

A nonlinear first stage does not generate consistent estimates in the second stage unless the nonlinear model happens to be exactly right, therefore the risk of misspecification is high (Angrist and Krueger, 2001). The coefficient in an LPM model also has a straight-forward interpretation: The π implies an increase in the probability of taking a college degree. For example a π of 0,01would imply a 1 percent point increase in the probability of taking a degree for individuals who had access to a college in their home region at the age of 19/25.

Using fixed effects and cohort dummies allow for a comparison between the same cohort, living in region with and without a college, and individuals from the same region belonging to pre- and post-reform cohorts. As fixed effects remove the effect of time-invariant factors, we are not estimating the effect for regions that had college the whole period (such as the big cities). If we assume that the error term conditional on $Cohort_i$ and $Region_i$ is uncorrelated with *Access to Colleg*e, the parameter π can be interpreted in a causal fashion. In other words there should be no other factors determining both access to college and college degree. Whether this is actually the case is however far from obvious. I will address instrument exogeneity and the common trend assumption in section 6.2 .

5.2 Second stage: Assessing returns to college degrees

Did the college expansion have any effect on income for those that were "pushed" into a field of education because a college was established in their home region? To explore this question a two stage least square (2SLS) regression is used, instrumenting college attainment with "Access to college" from in the first stage. The following model is used to see if the exogenous variation in access to college had any effect on income through attainment of nursing, engineering, teaching or a business degree:

$$LogWage = \beta_{iv}\widehat{Degree}_i + \textstyle\sum_{1932}^{1972}\lambda Cohort_i + \sum \gamma Region_g + \theta Controls_i + u_{ig} \qquad (2)$$

The logarithm of income is the dependent variable, measured as average income for individuals when they are 35-40 years. The independent variable, attainment of *college degree* is suspected to be correlated with some unknown factors, such as ability, that both affect degree attainment and earnings. We are therefore interested in the causal effect of these college degrees on earnings. Therefore college degree is instrumented by access to college at the age of 25, the fitted values $\widehat{Degree}_i$from the first stage. We are looking at the effect on earnings for those who went to college because a college was established in their home region compared to those that did not have a college nearby. An IV-estimation is biased if an omitted relevant variable is correlated either with the included non-endogenous explanatory variable or the instrumental variable. Therefore, as before, fixed region effects and a full set of birth year dummies are added. Also dummies for woman and parental education are added in one regression. The error term is as before clustered at the regional level[26].

The incomes are not adjusted for inflation. This is because the difference in incomes over time is accounted for when including cohort specific dummies. Still, when using nominal incomes, the differences in incomes between younger cohorts are given more weight because they are of a greater absolute value. Using log incomes takes away much of this effect. In many countries the college-premium has increased over time. This could be a potential threat to this model as this would also give more weight to younger cohorts. However, analysis of Norwegian data shows that the returns to education in Norway has remained stable over time (Hægeland et al., 1999 and Barth et al., 1999).

As the period we are looking at is very particular for the female share of the labour force, in the sense that women increased their labour market participation dramatically. This could potentially be affected by the establishment of education institutions offering female dominated fields of study. Therefore a separate regression is performed looking at the establishment of a nursing, teaching and regional college and female labour force participation .

[26]The correct clustered standard errors using fixed effects in a 2SLS regression are obtained using Stata's xtivreg2 program.

6 First Stage Results : Is the Establishment of Colleges a Good Instrumental Variable?

A good instrument is should be correlated with the endogenous regressor for reasons that the researcher can verify and explain, but should have no effect on the outcome variable through any other channel (Angrist and Krueger, 2001). Therefore a good instrument must be based on detailed knowledge on institutions and decision mechanisms. Previously, background knowledge is used to argue that the college expansion offers exogenous variation and is therefore not correlated with the outcome variable. A fixed effect regression relies on the assumption that without the intervention, the treated and untreated districts would have had the same trend. An attempt to graphically see if this is the case follows after the regression results.

6.1 First stage results

Table 2 shows that the baseline regressions (i-iv) are significant for all professions and do not change much when adding controls for parental education and gender. I find no significant effect regressing access to regional college on higher education (v). In the latter a dummy equal to one if an individual has higher education is regressed on access to a regional college. In the model used here I am not able to use the establishment of regional colleges as an instrument for higher education.

Not surprisingly, the results for nursing and engineering are only significant for women and men respectively. Since women tended to take more teaching degrees, the effect is slightly larger in size for women, but interestingly, more significant for men. Overall business had the most significant effect. This is not surprising as this degree was introduced with the regional colleges. When using 46 regions, the coefficients decrease in sign or remain the same, but are still significant. This suggests that with 90 regions we are identifying more local effects.

The overall effect conceals great heterogeneity based on social background. The effect is largest in size for those whose parents had short higher education. These results confirms

Aamodt (1982) who finds that students tended to take education at the same level as their parents. Teaching is different in the sense that it only yields significant results for the group with high school educated parents. Only nursing and business give significant results for the group whose parents had completed no more than elementary school. As we lack information on parental education for about 20 % of the population studied here, the results could be misleading. As seen in Figure 1, more than 50% had basic school in 1970, while in the data about 36 % have parents with this level as their highest education. This suggests that this group could include many with missing social background. In addition, from the summary statistics we know that the group with missing parental education level undertook the specific educational fields with similar frequencies as the elementary-school group.

Adding dummies for parental education level and gender (3) does not change the results substantially. A possibility could be to also control for educational level in order to compare "more similar" individuals. The results are shown in Appendix B. The size of the estimates drop slightly, but they continue to be significant, at least at the 5 % level. However, educational level can be a "bad control"[27], because educational level may itself be an outcome of the instrument: People obtained a higher education level because they were affected by a college establishment.

In Table 2 only the coefficients on access to college are shown because this is the effect that the first stage is addressing. Each column represents the different independent variables: access to nursing, engineering, teaching or regional colleges. The rows represent the different specifications. Therefore each coefficient represents one separate regression.

[27]A term discussed in Angrist and Pischke, "Mostly Harmless Econometrics", p 64-68.

Table 2: First stage results: Effect of college establishment on degree attainment

		(i) **Nurse**	(ii) **Engineer**	(iii) **Teacher**	(iv) **Business**	(v) **High ed.**
(1)	**Baseline**	0,0084***	0,0075***	0,017**	0,01***	0,0066
	(N=2309176)	(3,47)	(2,85)	(2,48)	(4,82)	(0,91)
(2)	**Region 46**	0,084***	0,0074**	0,014**	0,0079***	0,00012
	(N=2309176)	(2,70)	(2.62)	(2,02)	(3,20)	(0,02)
(3)	**With Controls**	0,0073***	0,0074**	0,017**	0,011***	0,008
	(N=1852402)	(3,25)	(2,56)	(2,62)	(4,57)	(1,01)
(4)	**Excl. Oslo**	0,0051***	0,0041*	0,015*	0,0084***	0,0073
	(N=1746516)	(2,72)	(1,84)	(1,82)	(4,01)	(1.05)
	By gender:					
(5)	**Women**	0,016***	0,0014	0,021**	0,0089***	0,012
	(N=1124659)	(3,56)	(1,43)	(2,04)	(4,11)	(1,16)
(6)	**Men**	0,0009	0,014**	0,012***	0,012***	0,0019
	(N=1184517)	(1,42)	(2,53)	(3,72)	(4,54)	(0,30)
	By parental ed.					
(7)	**H.E: 4 years+**	0,012**	-0,0013	-0,11	0,016***	-0,014
	(N=85017)	(2,40)	(-0,12)	(-0,95)	(3,14)	(-0,66)
(8)	**H.E 1-4 years**	0,017***	0,014***	0,0053	0,018***	0,011
	(N=184624)	(3,18)	(4,53)	(0,40)	(4,34)	(0,83)
(9)	**High School**	0,0069**	0,0087**	0,015***	0,013***	0,013
	(N=900960)	(2,41)	(6,86)	(3,91)	(5,11)	(1,05)
(10)	**Elementary School**	0,0046**	0,0051	0,017	0,0039***	0,0089
	(N=684813)	(2,03)	(1,48)	(1,44)	(2,85)	(0,84)

Notes: t-statistic in parenthesis. The standard errors are clustered at the region level. Region fixed effects and cohort dummies are included in all regressions. All columns present the coefficient on "Access to College" from a regression of attainment of specific college degree on access to college in home region when 25 years old. (1) is the baseline regression. In (2) 46 regions are used, in (3) parental education and gender dummies are added, in (4) Oslo region is excluded, in (5-6) baseline regression is done separately by gender, and in (7-10) by parental education level.
*p<0,10, **p<0,05, ***p<0,01.

6.2 Methodological Concerns

Instrument exogeneity

A main concern when using an instrumental variable is exogeneity. This cannot be tested empirically, so we have to look to historical sources to approach an answer. Was the location of these institutions of higher education exogenous? When it comes to the regional colleges, this depended on the different locations. While some of the first (like Stavanger and Kristiansand) were results of substantial local lobby activity, the last were mostly located so that every county had its own college. For the nursing and engineering colleges, this also seems to be the tendency, except that the first ones were established earlier, some even in the 1800's. Thus, the biggest and most important centers already had colleges before this expansion. Therefore the cities that got new colleges in this round were probably less different from the ones that did not get. Also, the fact that the expansion was based on the an idea that every education region should have a college suggests that the locations were chosen mostly to meet some geographical distribution criteria. Hence, the reforms proceeded "filling in" colleges in the different regions. .

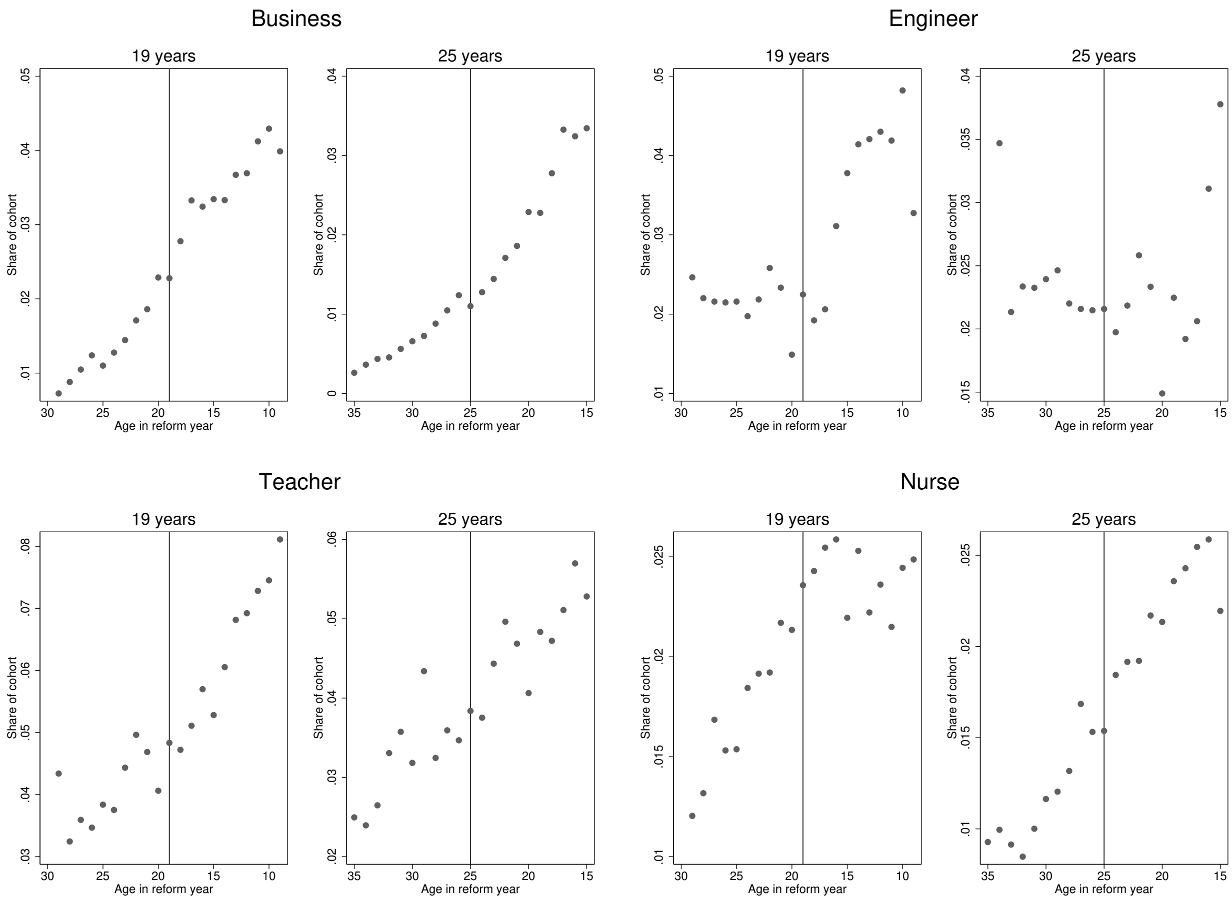
Still, assuming that the location of colleges was exogenous in the strictest sense would be rather speculative. Nevertheless, given the considerations above, it seems reasonable to assume that there was at least some exogeneity in the timing of the location of colleges, and by that the exogeneity requirement is satisfied. Therefore the analysis will further look at what effect the reform had and interpret the results with caution. According to historical sources (Johnsen, 1999), the location of teacher colleges established in the 1950's and 1960's were mostly a response to local demands. Therefore teaching might fail to meet the exclusion restriction and is therefore not a valid instrument.

Valid fixed effects assumptions?

The fixed effect model controls for unobservable factors across regions and across time. Essential to causal inference in a fixed effects model is the assumption that the treated units experienced a different trend because they were treated and that there are no other time-variant unobservable factors correlated with college attainment. We assume that the treated regions would have experienced the same trend as the untreated in the absence of treatment. This is not testable since we cannot observe the counterfactual: What would have happened in a region without a college if a college was established there? A possible convincing argument is to see the trend graphically in treated regions. Was there a large increase close to the year of establishment? In order to investigate this, I try to look for "jumps" in the proportion of a cohort taking a particular college degree.

Figure 3 shows the proportion of cohorts in regions where a college was established taking the degrees of interest. In Appendix B.3, the share of cohorts taking specific college degrees is shown individually for every region where a college was established. It is not clear here whether there is a "jump" at the age of 25 or 19. However, we do not expect a clear discontinuous jump at any particular age. As previously explained, the reform affected many cohorts because the starting age of education varied. Nursing for example, before it was upgraded to college level, was available also for 16-year-olds. In addition it is reasonable to believe that the colleges increased their capacity after the establishment. The number of students depended on how many the colleges could enroll. The business degree was for example very popular. In 1980 many colleges enrolled less than 20 % of the applicants to this degree (Brattvåg, 1981) .

Figure 3: Share of cohort taking degrees in region with new college



Note: The figures show the share of cohorts born between 1932-1972 from regions where specific colleges were established, taking the specified degrees. The x-axis measures the age of a cohort in the year of a college establishment.

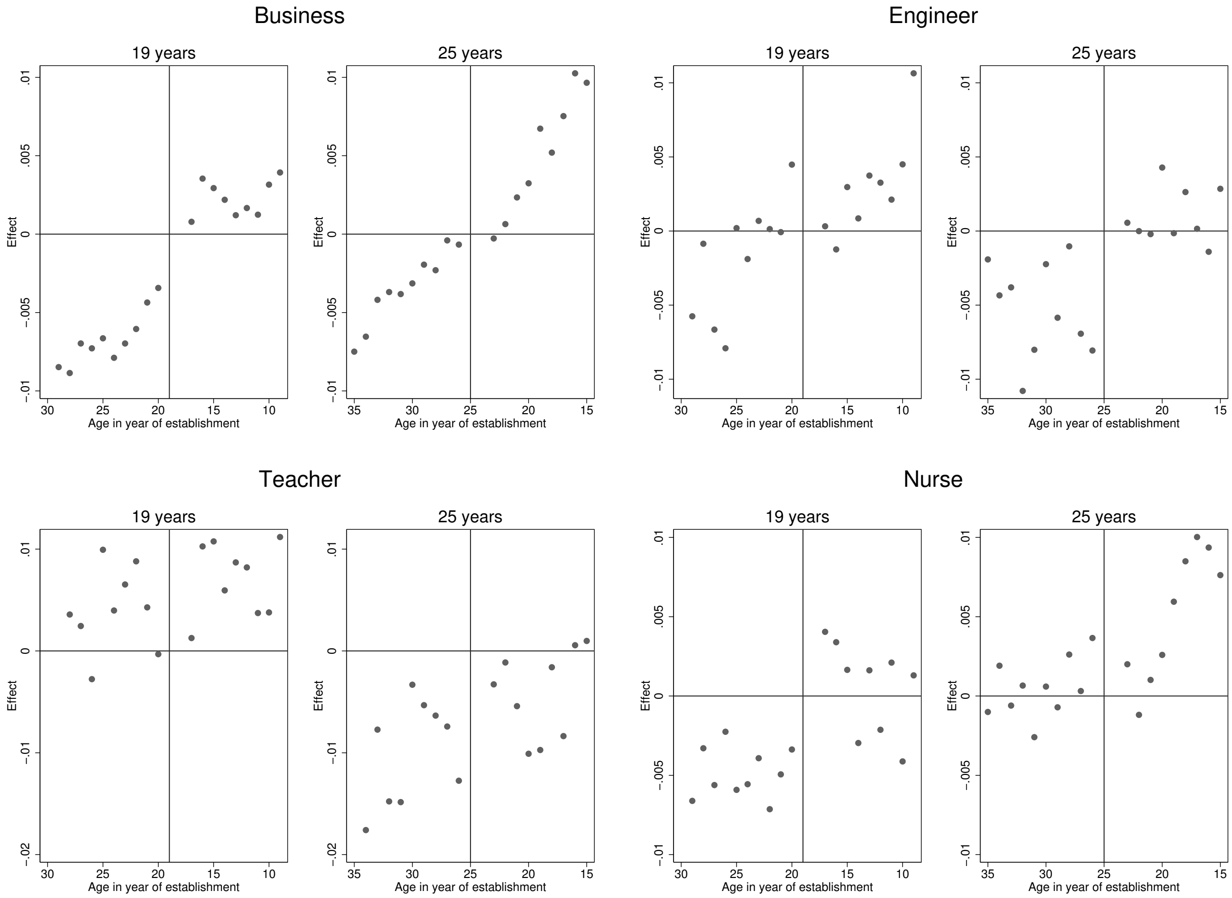
A possible concern with Figure 3 is that the cohort specific trend is not taken into account. Figure 2 showed that there was a sharp increase, in particular in nurses and engineers for the cohorts born in the middle 1950's. To overcome this issue the cohort specific effect must be isolated from the effect of college establishment. The following regression is performed in order to investigate whether timing of the reform is consistent with the true effect:

$$CollegeDegree_i = \alpha + \sum_{-15}^{15} \beta ReformCohort_i + \sum_{1932}^{1972} \lambda Cohort_i + \sum \gamma Region_g + \varepsilon_{ig} \qquad (3)$$

Here *Reform Cohort* is a set of 30 dummy variables from -15 to 15 where 0 indicates the cohort aged 19 or 25 years in the year of college establishment. 15 is chosen to capture variation in a large time-span before and after the reform. The value -15 is given to those born 15 years or more before the reform cohort or those from regions where colleges were never established. 15 is given to those born more than 15 years later than the reform cohort, thus this includes regions where there had "always" been a college. The 0-dummy is dropped so that the coefficients are the values relative to the reform cohort. Again region fixed effects and full set of cohort dummies are used. College degree is, as before, a dummy variable equal to one for individuals with the specified degree. The error term is again clustered at the regional level. The x-axis measures the age in the year of college establishment and the effect is measured relative to the cohort who were 19 or 25 years old that year. Figure 4 shows the coefficients from (3) using cohorts aged 19 and 25 years the year of college establishment respectively.

Having in mind that defining the reform cohort is difficult, these figures do suggest positive effects: The probability of taking the specified degrees increases with the reform for engineering, nursing and business. However, for teachers the trend is not so clear, possibly because as mentioned previously, local demand was an important factor in deciding the location of teacher colleges. A possible reason for why the increase was less important is that people from these regions were already pursuing teaching degrees.

Figure 4: Plot of coefficients - Effect of college establishment



Note: The figures show the coefficients on year born relative to reform cohort, from a regression of college degree on cohort dummies and distance to reform cohort dummies with region fixed effects. The x axis measures the age of a cohort in the year of a college establishment.

The Local average treatment effect- Who are the compliers?

Using a potential outcome framework, the local average treatment effect (LATE) is an average treatment effect for a subset of the population: Units that receive treatment if and only if they are induced by the exogenous instrument (Angrist and Pischke, 2008). These are individuals who undertake a college degree because a college was established in their home region and are called compliers. Although we cannot identify a complier, we can identify their characteristics.

Who were affected by the college expansion? Going back to previous findings, what we might expect is ambiguous: On the one hand, individuals whose parents had higher education would probably have taken higher education anyway by moving to another city. This suggests that having a college nearby should mostly affect those who would not have entered higher education otherwise. On the other hand, empirical studies previously referred to show that it was mostly people whose parents had high school education or more that entered higher education. Even though it was possible to leave for a big city, also children from these backgrounds might have preferred to take education closer to home. Thus the effect we might see is not the effect of taking higher education or not, but rather that the location of a college affected the choice of field. The fact that the effect is the smallest for groups who have the lowest predicted education, does not seem strange given that they had a smaller probability of entering higher education. Therefore rather than being induced to enter higher education, many were "pushed" into specific fields of study. Thus the "typical" compliers was an individual with relatively high educated parents, but who either because of abilities or taste, preferred education closer to home or was not able to enter a university.

In addition we know that the choice of field of study is highly reliant on preferences that are not so easily captured in a model. As previously mentioned, the data from Kirkebøen and co-authors (2016) tells a story of preferences[28] for areas of study: the establishment of a nursing college could push someone preferring nursing, teaching or related fields, to pursue a nursing degree, while a person preferring technical subjects was not affected at all. The results do suggest that this was the case: the establishment of colleges do not seem to have changed the gender pattern in nursing and engineering. However, there is some evidence that men were induced to enter a teacher college. In other words, a complier seems to be someone that had preferences to enter the specific field of study.

[28]Preferences that can be influenced both by individual characteristics such as ability and interest, but also social norms and expectations.

7 Second Stage Results: Returns to College Degrees?

The second stage regressions are made similar to the first stage: 1) using OLS, 2) the reduced form (regressing income on the college dummy from the first stage), 3) 2SLS, 4)Controlling for gender and parental education, 5) excluding Oslo area 6-7) by gender and 8-11) by social background. The results are shown in Table 3.

This is a log linear model. Therefore the interpretation of β_{iv}, is that a college degree increases the logarithm of income with β_{iv} units or income with a factor $e^{\beta \text{iv}}$. The OLS coefficient is positive for all, except for nursing. Thus, by controlling for birth year and region, individuals with these degrees earn relatively more. The results from the instrumental variable regression has the opposite result; only nursing has a positive and significant effect.

However, the coefficient on nursing, 4.45, is surprisingly large, and arguably too large to be credible. We can look to the reduced form regression for an explanation: The reduced form coefficient is larger than the first stage coefficient. The 2SLS coefficient is roughly the reduced form coefficient divided by the first stage coefficient. Here, that means: $\beta_{iv} = cov(Income, College)/cov(Degree, College)$. When the numerator is larger than the denominator we get a large coefficient. Either the nursing colleges had spillover effects that led to income gains, or these regions had some other characteristics that contributed to the rise in income. Therefore it seems plausible that a nursing degree led to higher income, but this is probably not the whole story. Thus, in order estimate the effects of college degrees on income, the establishment of nursing colleges might not be a good instrument since there seems to be something else, not controlled for, that affected both income and nursing college. For example, if a nursing college was established together with a hospital that brought with it other well pad jobs, this could be an explanation.

The reduced form estimate for nursing is larger in size and more significant for women than for men. The reduced form estimated separately by gender is shown in Appendix C.2. So, either the nursing colleges were established in places were women had more earning potential or the nursing colleges raised female wages. Although not significant, the reduced form estimates by gender show the same pattern for the other professions. Male-dominated engineering and business have larger reduced form estimates for men, while female-dominated teaching has a positive estimate for women and a negative for men. Having in mind the discussion on randomness in timing of college establishment, these results suggest the possibility that the college establishments did affect gender wage disparities.

Engineering colleges appear to have had a very small and not significant payoff. It is important to notice that the first stage for engineering is weak, slightly lower than the rule-of-thumb F-value of 10. With a weak instrument, tests of significance have incorrect size and confidence intervals are wrong. If the first stage F-value is weak , the 2SLS estimate will be biased towards the OLS estimate. In case of a weak instrument, a safer option is to look at the reduced form, which in this case is close to zero. Angrist and Krueger (2001) note that if you can't see the causal relation of interest in the reduced form, it's probably not there. Engineering had the strongest first stage for those whose parents had completed high school or short higher education. The effect on income is positive for the latter, while close to zero for the former group, though neither is statistically significant.

Business on the other hand, has a stronger first stage, but also little significant effects on income. When assessing the effect separately by parental education level, so that individuals are compared to others with similar background, the results suggest heterogeneous outcomes. The coefficient is positive and significant at the 10 % level for the group with college-level educated parents while it is negative for those whose parents had less education.

Teaching had the strongest first stage for men and for people with high-school graduated parents. None of these have any significant effect on income, but negative in sign for men and slightly positive for group 3. Having in mind that Figure 4 reveals no visible jump in the share of cohorts close to 25 the year of college establishment with teaching degrees, the second stage results should be interpreted with caution. Figure 5 shows the coefficients[29] from Table 3 in order to illustrate their precision (or lack thereof) and their distance from zero.

[29] Using Stata program "Coefplot".

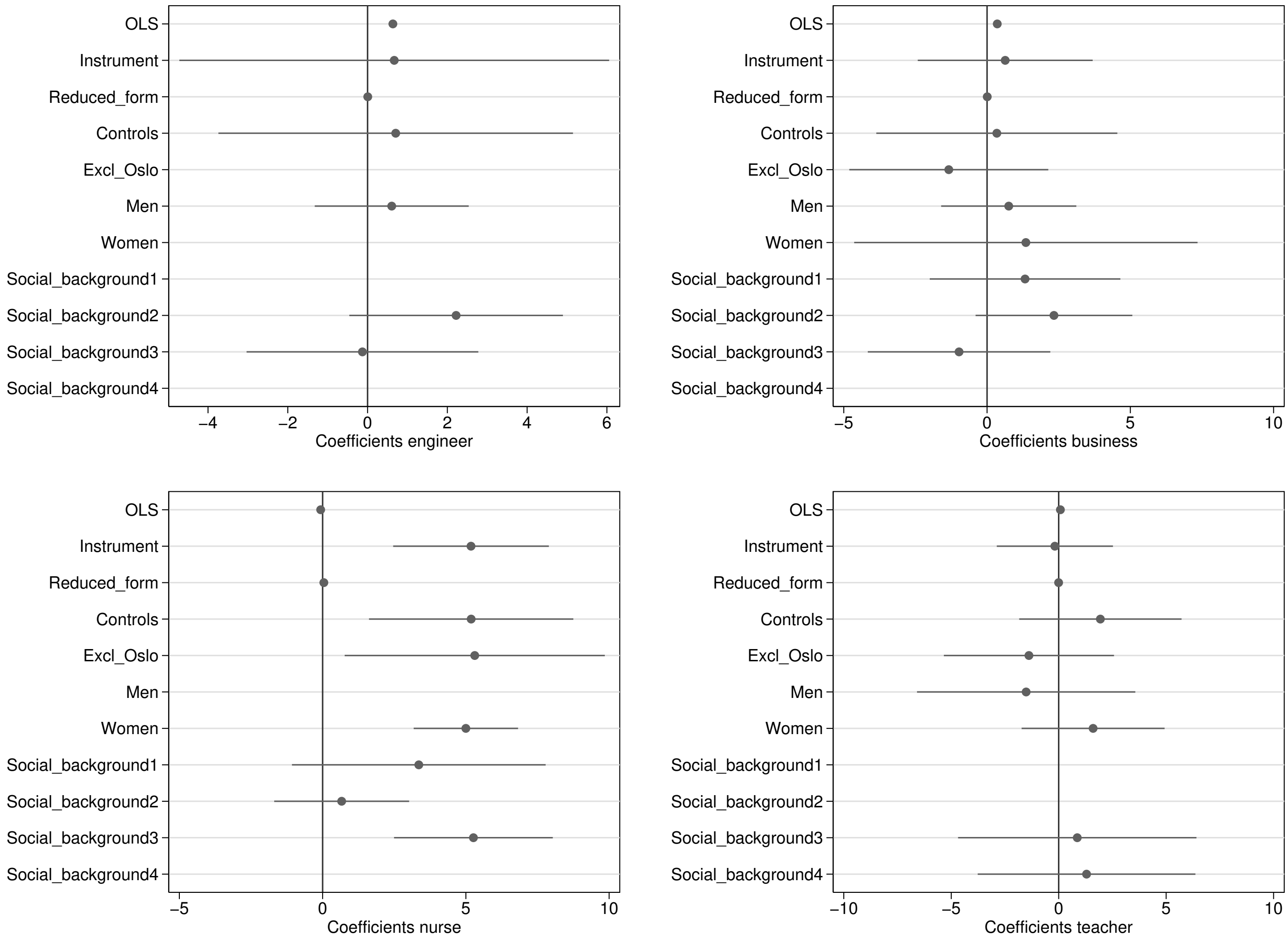
Table 3: Two Stage Least Square Regressions: The effect of college degree on wage

		(i) **Nurse**	(ii) **Engineer**	(iii) **Teacher**	(iv) **Business**
(1)	**OLS**	-0,046***	0,60***	0,094***	0,37***
	(N=2150940)	(-2,78)	(92,85)	(5,72)	(40,31)
(2)	**Reduced form**	0,038*	-0,0011	-0,0021	0,0051
	(N=2150940)	(1,97)	(-0,05)	(-0,11)	(0,28)
(2)	**2SLS**	4,45***	-0,15	-0,093	0,49
	(N=2150940)	(2,94)	(-0,05)	(-0.08)	(0,29)
	F-stat first stage	12,21	7,89	6,41	23,35
(4)	**With Controls**	4,72***	0,88	-0,31	0,65
	(N=1743826)	(2,64)	(0,42)	(-0,20)	(0,31)
	F-stat first stage	10,20	6,29	6,85	21,12
(5)	**Excl. Oslo**	4,69*	-4,13	-1,25	-1,18
	(N=1675680)	(1,86)	(-0,60)	(-0,48)	(-0,59)
	F-stat first stage	7,44	2,53	3,80	16,21
	By Gender				
(6)	**Women**	3,44***	-0,030	1,73	0,35
	(N=1019017)	(4,46)	(0,00)	(1,35)	(0,12)
	F-stat first stage	12,38	3,26	4,51	16,98
(7)	**Men**	32,9	0,81	-1,25	1,00
	(N=1131923)	(1,47)	(0,81)	(-0,48)	(0,78)
	F-stat first stage	3,72	5,54	13,71	20,71
	By Parental Ed.				
(8)	**H.E: 4 years+**	5,95*	86,7	9,34	2,22
	(N=79036)	(1,71)	(0,02)	(0,56)	(1,35)
	F-stat first stage	5,40	0,00	0,75	10,14
(9)	**H.E 1-4 years**	1,64	0,71	1,55	2,77*
	(N=174697)	(1,14)	(0,70)	(0,50)	(1,67)
	F-stat first stage	10,28	20,26	0,13	19,59
(10)	**High School**	6,34***	0,03	0,32	-0,64
	(N=851781)	(5,47)	(0,49)	(0,13)	(-0.39)
	F-stat first stage	5,21	8,32	14,99	26,35
(11)	**Elementary School**	6,75	0,040	1,31	-0,33
	(N=635422)	(1,37)	0,007	(0,53)	(-0,04)
	F-stat first stage	5,17	2,64	1,90	8,24

Notes: t-statistic in parenthesis. Region fixed effects and cohort dummies are included in all regressions. The standard errors are clustered at the region level. In (1) wage is regressed on having the specified degrees, (2) is the reduced form showing the coefficient from a regression of wage on access to college, (3) is the baseline IV-regression where college degree is instrumented with access to college when 25 years old, in (4) the Oslo region is excluded, in (5-6) the baseline IV regression is done by gender and in (7-10) by parental education level.

*p<0,10, **p<0,05, ***p<0,01

Figure 5: Plotted coefficients from 2SLS regressions: Effect of college degree on wage*



*Coefficients whose confidence interval exceeds ≥|20| are dropped from the plot.

Female labour supply

If the establishment of colleges did affect gender wage disparities, a possible mechanism is that a college establishment caused women that otherwise would not have worked, to enter the labour market. The analysis of female labour supply is very similar to the 2SLS regression explained previously. Instead of wage as the dependent variable, a dummy variable equal to 1 if an individual worked between 35-40 is used. Working is defined as having an income above twice the pension base rate [30](e.g. in 2013 this corresponds to 170 000 NOK). Thus individuals whose average wage is less than twice the average base rate during the years when they are 35-40 years are considered to be non-working. Any income below this is not likely to come from full labour force participation.

Table 4: Effect of nursing, business and teaching college establishment on female labour supply

		(i) **Nurse**		(ii) **Business**		(iii) **Teacher**	
		OLS	**2SLS**	**OLS**	**2SLS**	**OLS**	**2SLS**
(1)	**College**	0,042***		-0,002		0,023**	
	(N=1122345)	(3,68)		(-0,20)		(2,24)	
(2)	**Degree**		2,91***		-0,24		1,05*
	(N=1122345)		(3,53)		(-0,19)		(1,80)

Notes: t-statistic in parenthesis. Standard errors are clustered at the region level. Region fixed effects and cohort dummies are included in all regressions. A dummy for working is the dependent variable and is equal to 1 if an individual's average income when 35-40 years old is above twice the pension rate, and 0 otherwise. College (1) presents the reduced form: An OLS regression on access to college. Degree (2) presents a 2SLS regression where college degree is instrumented with access to respectively nursing, business and teaching college when 25 years old.
*p<0,10, **p<0,05, ***p<0,01

Table 4 shows that nursing and teaching degree, when instrumented with access to college has a positive effect on employment. It seems strange that the coefficient is larger than one, showing the disadvantage of the linear probability model. The reason for the large coefficient in case of nursing is probably the same as previously discussed, that the reduced form is larger than the first stage coefficient and there seems to be some other factors influencing work in regions that have a nursing college. The same argument applies to the coefficient on teaching, though this one is less significant and has a weak first stage for women. Nevertheless,

[30]A base rate used as the basis for calculating the Norwegian state pension. The pension base rate is adjusted annually. I also did the regression using 1,5 G as the threshold and got very similar results.

the idea that the establishment of nursing and possibly teaching colleges, offering women a college degree, caused more women to work, seems plausible. The business degree, which was male-dominated, on the other hand, seems to have had no effect on female labour supply.

Why do we see little returns in the college degrees?

As in many other attempts at estimating returns to education using IV-methods, the results found are relatively imprecise and should therefore be interpreted with caution. An instrumental variable estimator identifies a Local Average Treatment Effect (LATE), which measures the return to college for individuals induced to go to school by the change in the instrument[31]. It is important to note that it is therefore unlikely that we are estimating the average return to these degrees.

There are many possible explanations why nursing is the only field that shows some positive significant effect on income, although also this result must be interpreted with caution. Research repeatedly has shown that women have a higher return to education than men (Harmon et al, 2000). A nursing degree gave women higher monetary returns than the alternatives, which for example were not to work or take a lower paid service sector job. Other possible explanations are some characteristics of the nursing profession, such as a standardized degree and wages that are subject to centralized bargaining, which could contribute to less wage disparities between nurses. Therefore the "marginal" nurse should have a wage close the average nurse, although the "marginal" nurse might have had less opportunities in other parts of the labour market. Since it is the group with high school educated parents that drives the results, this indicates that for this group nursing gave more returns than the alternatives.

The results from the first stage regressions have shown that those induced to enter the new colleges were not necessarily the less privileged groups. Card (1995), finds a high return to

[31] In a potential outcome framework, in order to estimate the LATE, we must assume that there are no defiers, i.e that all are affected by the instrument in the same direction. Here it seems reasonable to assume that nobody chose *not* to undertake a college degree as a result of college establishment

college for those with the lowest expected length of education. In the case of the Norwegian college expansion, the group whose parents had little education, only elementary schooling, seems not to have been affected very much. This could be a possible reason small effects on earnings: those affected would in absence of a nearby college, have pursued another degree in another college. If this is the case the case the returns could have been higher if the low educated groups had been affected.

Another possible explanation is heterogeneity in returns to the college degrees; some of those affected by the instrument might have been people with low returns to education. Both teaching and business have the strongest first stage for those whose parents had completed high school. They were advancing one step in the education system compared to their parents. Being the first in a family to complete higher education might indicate that certain career and wage expectation can be different than for those from families where parents with high education can be "role models". In particular for business, the poor results for the lower social groups and positive results for higher social groups taking this degree might represent this issue. Business does not have a clear career path such as nursing. Higher social groups could have had other networks and therefore ended up with jobs in other sectors than the lower social groups. This can also be the case for the engineering degree that shows some sign of positive returns for higher social groups and no returns for others[32]. This does not support the idea that the returns to the college degrees would have been higher if the group with low educated parents had been affected to a greater extent.

The new colleges might have been of poor quality. Since many of these colleges were established based on an idea of equalizing across regions, and not because of local demand, they did not have the prerequisites to offer high-quality teaching. Johnsen (1999) documents how the newly established regional colleges had a "precarious" lack of qualified teaching personnel. Considering that the business degree was a result of decision-makers ideas of what a college should offer, and not local conditions, a possibility is that this degree was not in high demand in the local labor markets. An interesting extension of this work would be to

[32]The way the variable engineering is constructed can be decisive; it includes all degrees offered at engineering colleges. If higher social groups chose degrees with higher returns than others, this might explain the earning differences.

investigate in which jobs the business-graduated ended.

8 Concluding Remarks

This thesis contributes on one empirical and one methodological topic: The effects of the expansion of higher education and the use of an instrumental variable to study field of education. Did the expansion of higher education lead to a more equal distribution of education geographically and across social groups? Can this college reform be used as an instrumental variable?

A question implicitly raised both in the first and second stage regression is the success of the educational policies aiming at distributing higher education across Norwegian counties. The first stage shows that for the most part people nearby were affected. However, other than for nursing, I am not able to detect a positive effect on earnings. Is it therefore right to conclude that the educational policies failed? First, the cohorts included in the analysis lived in a specific period in a Norwegian context, a period of growth and expansion of the welfare state which perhaps gave gains to a large majority of the population. For example financial support to students increased and was made available to all and income inequality decreased. So in addition to the college expansion, other things happened that made the effect of college degrees and college proximity less decisive. Secondly, the scope of this thesis is also limited in that it attempts to estimate returns to some specific fields and cannot therefore necessarily be extended to other fields of education.

Finally, other than monetary gains, the colleges brought with them many jobs for people with high education outside the larger cities, and other non-private returns to education. The new colleges might have been of poor quality in the beginning, but the effect they have today has not been considered. There is no doubt that this reform has had lasting impacts on the geographical distribution of higher education institutions in Norway. The lesson learnt is that it is not obvious that the expansion of higher education has individual monetary gains, whether this is because the treated individuals have little gains from higher education (selection on gains) or that local resources are not sufficient to provide good quality education.

The next question raised is if the college expansion succeeded in increasing access to college across social groups. The findings in this thesis do not suggest that the expansion of higher education in the college sector did affect social mobility- neither in terms of income nor education. Overall, compared to the majority group who had only completed elementary schooling, the effect of the college establishment was larger for individuals whose parents had high school or 1-4 years of higher education. It was also the latter group that seems to have gained the most from a business and engineering degree. Thus the findings suggest that it was the least able from relatively high educated families rather than the most able from low educated families, that gained from this reform. An exception is seen for women whose parents had completed high school. They seem to have experienced monetary gains from a nursing degree. Although, the results suggest that also other things can have happened causing female wages to rise in these regions.

Even though theory suggests that decreasing distance to college should have a greater effect on the low educated groups, this is not necessarily the case empirically. Further research can use this reform to understand how expansion of educational opportunities in higher education affected intergenerational mobility and income inequality in Norway. On the one hand education was now more accessible for all and one would expect equalization, but on the other hand children from families with high education may have carried expectations from their families regarding college education and it was now easier for them to achieve this. They could study closer to home and entrance requirements were relaxed as there were more study places.

Previous usage of educational reforms as instrumental variables has mostly been concerned with the length of education. Wage inequalities are to a large extent also explained by field of study, which should therefore also be addressed. It is more challenging to find randomness in assignment to fields of education than in mandatory schooling at lower levels. We face an issue of selection-bias because people tend to choose fields in which they have a comparative advantage (Kirkebøen et al, 2016). Therefore we cannot estimate the average return of a college degree, but the return of the marginal degree taker.

This thesis has shown that validity of college expansion as an instrument for field of study depends on the specific field and the selection of college location. Its disadvantages is its weak correlation with degree attainment for some professions, difficulty in identifying a reform cohort, and the possibility that new education institutions more or less randomly located might have lacked prerequisites to offer high-quality education. Nevertheless, I believe that there is potential for further usage of this reform in research. As higher education was seen as a political means to geographic equalizing in Norway, it offers a rare exogenous variation in access to higher education across regions and time, and also field of study. Hence, the reform can be used to study the choice of field and not be restricted to length of education as other instrumental variable and natural experiment approaches are.

Understanding the choice of field of study is complex and has so far not received much attention in the economic literature. Educational reforms have been used to study intergenerational mobility with respect to length of education. Most studies find small or no effect of increases in mandatory schooling, while possibly a positive effect of higher education. The Norwegian college expansion can be used as an extension of this literature, by changing the focus from length to field of education.

The fields studied here are to some extent dominated by one gender, and therefore the complex dynamics with respect to the gendered labour market can be revealed. The results presented in this thesis give indications that the establishment of female-dominated fields is related to an increase in both female earnings and female labour supply. This confirms recent findings that a substantial portion of the rise in female hours is associated with the reallocation of labor from manufacturing into services (Olivetti and Petrongolo, 2016). The interplay between the rise in services and gender outcomes may have both supply and demand components and the expansion of colleges can possibly be further used to address these questions.

The college reform can also give new insights regarding the persisting gender differences in the choice of field of study. Hulum et al. (2016) document a significant relationship in the gender stereotypical nature of educational choices across generations using Danish register data. The college reform can contribute to study the causal nature of this relationship in a Norwegian context. For example, the children of those affected by the college expansion and their educational choices can be studied in order to understand the effect of mother's nursing education or father's engineering degree on offspring's educational choice.

References

Aaberge, R., A. B. Atkinson, and J. Modalsli (2016). On the measurement of long-run income inequality. empirical evidence from norway, 1875-2013. *Discussion Paper Statistics Norway*.

Aamodt, P. (1982). Utdanning og sosial bakgrunn [Education and Social Background]. *Samfunnsøkonomiske studier 51, Statistics Norway*.

Altonji, J. G., E. Blom, and C. Meghir (2012). Heterogeneity in human capital investments: High school curriculum, college major, and careers. *Annu. Rev. Econ. 4*(1), 185–223.

Angrist, J. and A. B. Krueger (2001). Instrumental variables and the search for identification: From supply and demand to natural experiements. *Journal of Economic Perspectives Volume15, pp 69-85*.

Angrist, J. and J.-S. Pischke (2008). *Mostly harmless econometrics: An empiricist's companion.* Princeton university press.

Barth, E. and M. Roed (1999). The return to human capital in norway: a review of the literature. *Returns to human capital in Europe: A literature review, ETLA-The Research Institute of the Finnish Economy, Helsinki*, 227–58.

Becker, G. S. (1964). *Investment in human capital: A theoretical analysis.* The University of Chicago Press.

Bhuller, M. S. (2009). Inndeling av norge i arbeidsmarkedsregioner. *notat2009/24 Statistics Norway*.

Black, S. E. and P. J. Devereux (2010). Recent developments in intergenerational mobility. *NBER Working Paper Series*.

Bostad (2007). Nou 2007:6 formål for framtida- formål for barnehagen og opplæringen, Norges offentlige utredninger, Norwegian Ministry of Education.

Bratti, M., D. Checchi, and G. De Blasio (2008). Does the expansion of higher education increase the equality of educational opportunities? evidence from italy. *LABOUR 22*, 53–88.

Brattvåg, T. (1981). *Søkning og opptak til distriktshøgskolene 1980*. Agder distriktshøgskole.

Cameron, A. C. and D. L. Miller (2015). A practitioners guide to cluster robust inference. *Journal of Human Resources 50*(2), 317–372.

Card, D. (1995). Using geographic variation in college proximity to estimate the return to schooling. *In L.N. Christofides, E.K. Grant, and R. Swidinsky, editors, Aspects of Labor Market Behaviour: Essays in Honour of John Vanderkamp Toronto: University of Toronto Press*.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica 69*(5), 1127–1160.

Clark, T. and Åsa Sohlman (2009). OECD reviews of tertiary education: Norway. Technical report, OECD. Available online at https://www.oecd.org/norway/37457548.pdf.

Griliches, Z. (1977). Estimating the returns to schooling: Some econometric problems. *Econometrica: Journal of the Econometric Society*, 1–22.

Hægeland, T., T. J. Klette, and K. G. Salvanes (1999). Declining returns to education in norway? comparing estimates across cohorts, sectors and over time. *The Scandinavian Journal of Economics 101*(4), 555–576.

Harmon, C., H. Oosterbeek, and I. Walker (2000). *The returns to education: a review of evidence, issues and deficiencies in the literature*. Centre for the Economics of Education, London School of Economics and Political Science.

Heckman, J. J., L. J. Lochner, and P. E. Todd (2006). Earnings functions, rates of return and treatment effects: The mincer equation and beyond. *Handbook of the Economics of Education 1*, 307–458.

Heineck, G. and R. Riphahn (2009). Intergenerational transmission of educational attainment in germany-the last five decades. *Journal of Economics and Statistics (Jahrbuecher fuer Nationaloekonomie und Statistik) 229*(1), 36–60.

Hernes, G. (1991). Stortingsmelding nr 40, fra visjon til virke, om høgre utdanning [white paper to the parliament about higher education].

Holmlund, H. (2008). Intergenerational mobility and assortative mating: Effects of an educational reform. *Centre for the Economics of Education, LSE*.

Humlum, M. K., A. B. Nandrup, and N. Smith (2016). Closing or reproducing the gender gap? parental transmission, gender and education choice. In *Determinants of Student Achievement and Education Choice*, pp. 57. Ph. D. thesis, Institut for Økonomi, Aarhus Universitet.

Johnsen, B. W. (1999). *Fra universitetsvisjon til høyskoleintegrasjon*. Ph. D. thesis, Kristiansand: Høyskoleforlaget.

Kirkeboen, L. J., E. Leuven, and M. Mogstad (2016). Field of study, earnings, and self-selection. *The Quarterly Journal of Economics 131*(3), 1057–1111.

Layard, R. and G. Psacharopoulos (1974). The screening hypothesis and the returns to education. *Journal of political economy 82*(5), 985–998.

Lindbekk, T. (1998). The education backlash hypothesis: The norwegian experience 1960-92. *Acta Sociologica 41*(2-3), 151–162.

Maurin, E. and S. McNally (2008). Vive la revolution! long-term educational returns of 1968 to the angry students. *Journal of Labor Economics 26*(1), 1–33.

Mincer, J. (1974). *Schooling, Experience, and Earnings. Human Behavior & Social Institutions No. 2.* New York: Columbia University Press.

Norges almenvitenskapelige forskningsråd (1979). *Høyere utdanningsinstitusjoner og forskningsinstitusjoner i Norge: En fylkesvis oversikt*. NAVF.

Norwegian Centre for Research Data. The Norwegian State Administration Database. Available online at http://www.nsd.uib.no/polsys/en/civilservice/.

Norwegian Ministry of Education (1964). St. meld nr. 69 om lærerutdanning [white paper to the parliament about the ededucation of teachers].

Norwegian Ministry of Education (1969). St. meld. nr. 63 om virksomheten i statens lånekasse for utdanning i regnskapsåret 1969 [white paper to the parliament about the central government's loans to students].

Norwegian Ministry of Education (1975). St. meld nr. 17 om den videre utbygging og organisering av høgre utdanning [white paper to the parliament about the further expansion of higher education].

Norwegian Nurse Association (2017). History of norwegian nurses. Accessed on : 27.02.2017, Available online at https://www.nsf.no/vis-artikkel/7981/17036/NSFs-historie.

Olivetti, C. and B. Petrongolo (2016). The evolution of gender gaps in industrialized countries. *Annual Review of Economics 8*, 405–434.

Ottoesen, K. (1969). St.prp. nr. 136 instilling om videreutdanning for artianere [recommendation on further education for high school graduates].

Pekkarinen, T., K. G. Salvanes, and M. Sarvimäki (2017). The evolution of social mobility: Norway during the twentieth century. *The Scandinavian Journal of Economics 119*(1), 5–33.

Pekkarinen, T., R. Uusitalo, and S. Kerr (2009). School tracking and intergenerational income mobility: Evidence from the finnish comprehensive school reform. *Journal of Public Economics 93*(7), 965–973.

Salvanes, K., K. Liu, and P. Carneiro (2014). The supply of skills and endogenous technical change: Evidence from a college expansion reform. *CESifo Area Conference on the Economics of Education*.

Salvenes, K. G. (2017). Inntektsforskjeller og sosial mobilitet i Norge [Income differences and social mobility in Norway]. In Barne, ungdoms og familiedirektoratet (Ed.), *Oppvekstrapporten 2017, Økte forskjeller- gjør det noe?*, pp. 64–76.

Statistics Norway (2007). Nasjonal utdanningsdatabase dokumentasjonsrapport. *Statistics Norway Reports 2007/54*.

Statistics Norway (2009). Lønnsgapet mellom kvinner og menn består [the wage gap between men and women persists]. Available online at https://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/lonnsgapet-mellom-kvinner-og-menn-bestaar.

Vassenden, K. (1987). Folke- og bolitellingene 1960, 1970 og 1980 - Dokumentasjon av de sammenlignbare filene. *Reports 87/2, Statistics Norway*.

World Bank (2012). Gender Equality and Development, World Development Report. Available online at https://siteresources.worldbank.org/INTWDR2012.

Appendix

A Higher Education Institutions by County

Establishment of Colleges[33] (or similar level institutions) in Norway by county with year of establishment in parenthesis:

Østfold

Engineer College: Sarpsborg(1965) ; Nurse College: Fredrikstad(1955); Regional College: Halden(1976); Teacher College: Halden(1963).

Akershus[34]

Engineer College: Ås(1897) ; Nurse College: Lørenskog(1895); Regional College: None; Teacher College: Stabekk(1909).

Oslo

Engineer College: 1873 ; Nurse College: 1868; University: 1811; Teacher College: 1875.

[33]In Oslo, Bergen, Trondheim and Stavanger, only the first college within a field is listed.
[34]Akershus belongs to the Oslo region according to the Labour Market Regions (Bhuller, 2009)

Hedmark

ENGINEER COLLEGE: none; NURSE COLLEGE: Elverum(1927); REGIONAL COLLEGE: Rena (1979); TEACHER COLLEGE: Hamar (1866).

Oppland

ENGINEER COLLEGE: Gjøvik(1966); NURSE COLLEGE: Gjøvik(1970); REGIONAL COLLEGE: Lillehammer(1971); TEACHER COLLEGE: none, but pedagogy was offered at the college in Lillehammer.

Buskerud

ENGINEER COLLEGE: Kongsberg(1970); NURSE COLLEGE: Drammen(1935); REGIONAL COLLEGE: None; TEACHER COLLEGE: Hønefoss(1985).

Vestfold

ENGINEER COLLEGE: Horten(1855); NURSE COLLEGE: Tønsberg(1918); REGIONAL COLLEGE: None; TEACHER COLLEGE: Tønsberg(1959).

Telemark

ENGINEER COLLEGE: Porsgrunn(1884) ; NURSE COLLEGE: Skien(1908); REGIONAL COLLEGE: Bø(1970); TEACHER COLLEGE: Notodden(1895) .

Aust-Agder[35]

Engineer College: Grimstad(1967) ; Nurse College: Arendal(1920); Regional College: Grimstad(1969); Teacher College: None.

Vest-Agder

Engineer College: None ; Nurse College: Kristiansand(1976); Regional College: Kristiansand(1969); Teacher College: Kristiansand(1839).

Rogaland

Engineer College: Stavanger(1878) ; Nurse College: Stavanger(1920) and Haugesund(1980); Regional College: Stavanger(1969); Teacher College: Stavanger (1955).

Hordaland

Engineer College: Bergen (1875); Nurse College: Bergen (1918) and Stord(1977); University: Bergen (1946) ; Teacher College: Bergen(1909) and Stord(1839).

Sogn og Fjordane

Engineer College: Førde(1988); Nurse College: Førde (1979); Regional College: Sogndal(1975) ; Teacher College: Sogndal(1973).

[35] Aust-Agder and Vest-Agder was together an education region (Johnsen, 1999)

Møre og Romsdal

Engineer College: Ålesund(1965); Nurse College: Molde(1958) and Ålesund(1974) ; Regional College: Molde and Volda(1969) ; Teacher College: Volda(1898).

Sør-Trøndelag

Engineer College: Trondheim(1900); Nurse College: Trondheim(1906) ; University: Trondheim(1969 from NTH established 1900) ; Teacher College: Trondheim(1922).

Nord-Trøndelag

Engineer College: None; Nurse College: Levanger(1940) and Namsos(1939) ; Regional College: Steinkjer(1980) ; Teacher College: Levanger(1938).

Nordland

Engineer College: Narvik(1955); Nurse College: Bodø(1930); Regional College: Bodø(1970) ; Teacher College: Nesna(1828) and Bodø(1952).

Troms

Engineer College: Tromsø(1987, offered at the University); Nurse College: Tromsø(1940) and Harstad(1982); University: Tromsø(1972); Teacher College: Tromsø(1824).

Finnmark

Engineer College: None; Nurse College: Hammerfest(1960); Regional College: Alta(1977); Teacher College: Alta(1974).

B Data

B.1 Course codes / Education codes

Course codes from NUDB are used for people registered with educations after 1970. These codes are found in the NUDB register. For individuals who are found in the census, but not in NUDB, their highest education at 3 code level registered in the 1970 census is used[36]. The following codes are used to identify the targeted fields of study:

Nurse: Code in NUDB: 661104, 3-year nursing school. All degrees are coded with the level that they have today (thus nursing is level 6 even though it has been at level 4). If not in NUDB, but in 1970 census: 471, 571 are used. Most have code 471 ("Nursing subjects") since this education was at high school level.

Engineer: The college degree variable in the case of engineering consists of engineer educations that were offered at Engineer Colleges. These are identified in the NUS-register as some degrees are are specifically labeled engineer college. These are 14 different degrees of which 5 had more than 10 000 students. These are: Electrical Engineer, Mechanical engineer, Construction Engineer , Computer Engineer [Databehandling] and Chemical Engineer. Codes in NUDB marked with Engineering College: 65501, 655299, 655207, 651301, 657105, 659921, 657104, 641115, 652205, 65530 659910 655102 652209 655105. From 1970 census the following codes are used: 555(Mechanical engineer), 555 (Electrical engineer), 557(Chemical engineer) and 558 (Construction engineer) . In addition people who according to the 1960-census had education from a engineering college (called technical school at this time) are included. The technical school had the institutional code 11.

Teacher: Code in NUDB: all codes starting with 62- are included (with an exception of education in pedagogy which was usually offered at universities). Many teachers

[36] These codes are documented by Statistics Norway here:https://www.ssb.no/a/histstat/rapp/rapp_198702.pdf

will have something more than a basic teaching degree, but as the data set is collapsed to one observation per individual this does not matter. From 1970 census 532, 533 and 632, 633 are used.

BUSINESS: Code in NUDB: 641115, a 2 year degree in business and administration ["Høgskolekandidat i økonomi og administrasjon"].

B.2 Maps of regional divisions

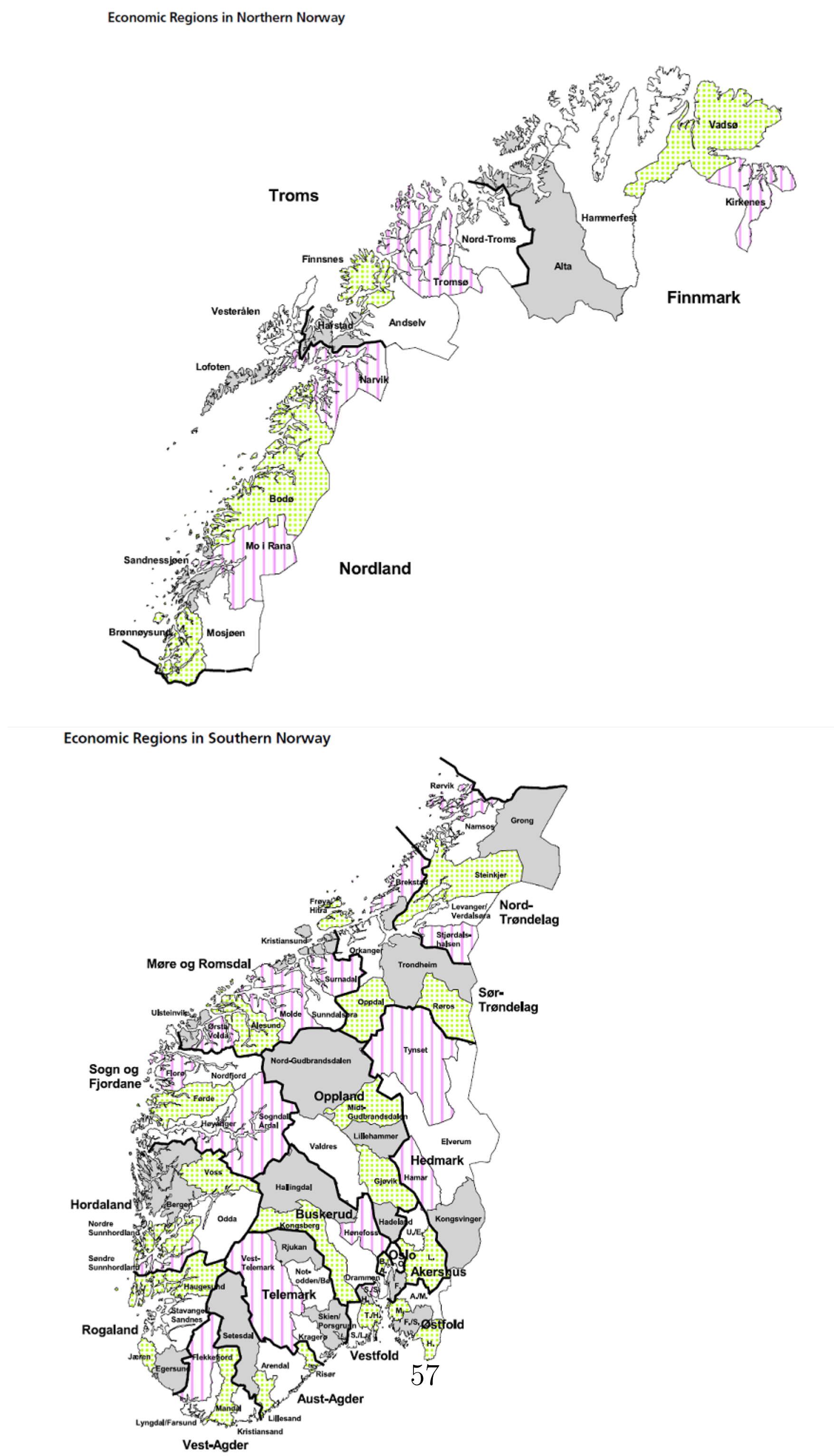


Figure B.1: Map of 90 regions, Classifications of Norwegian Regions
(Maps from Statistics Norway, 2000)

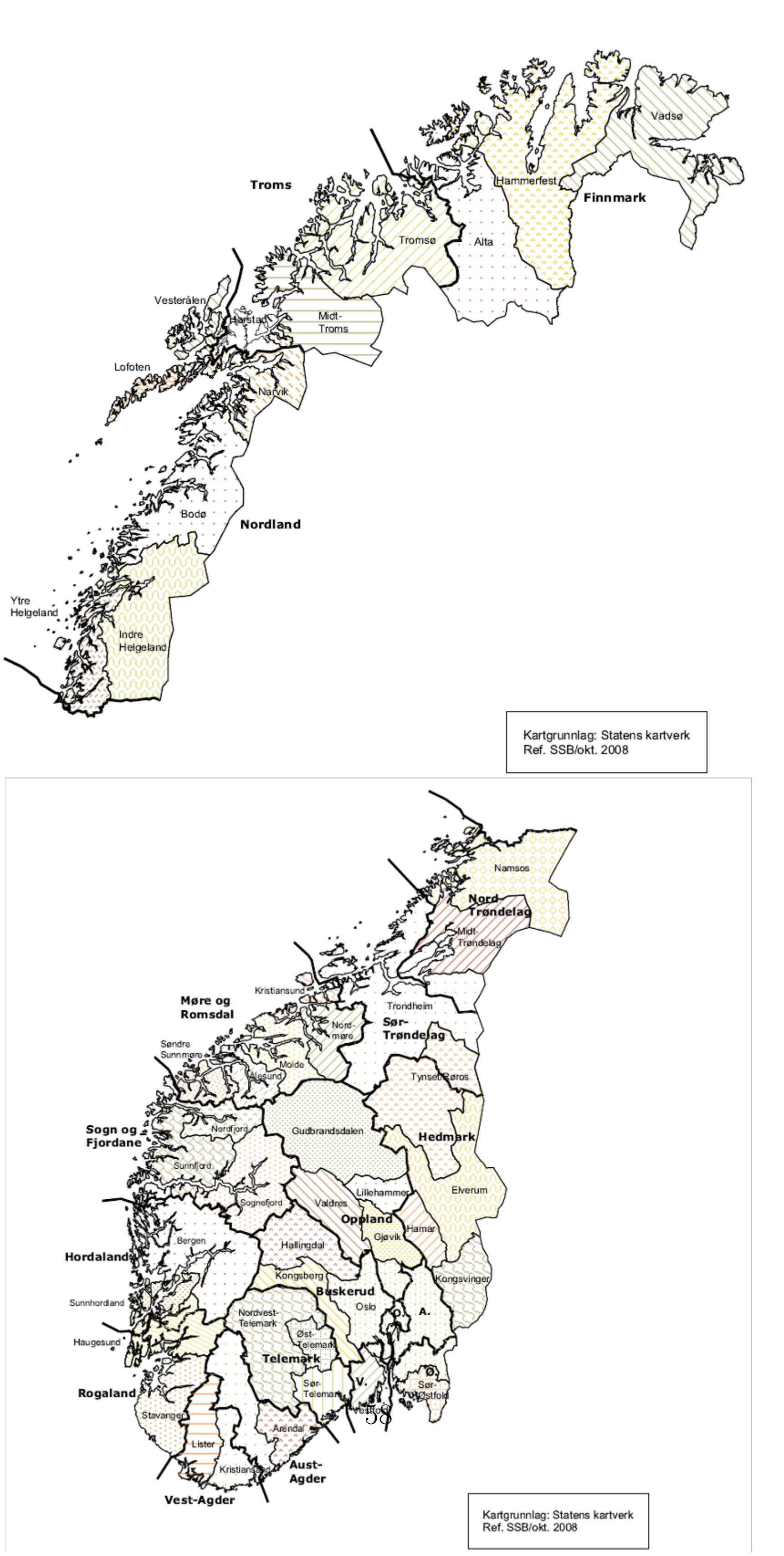
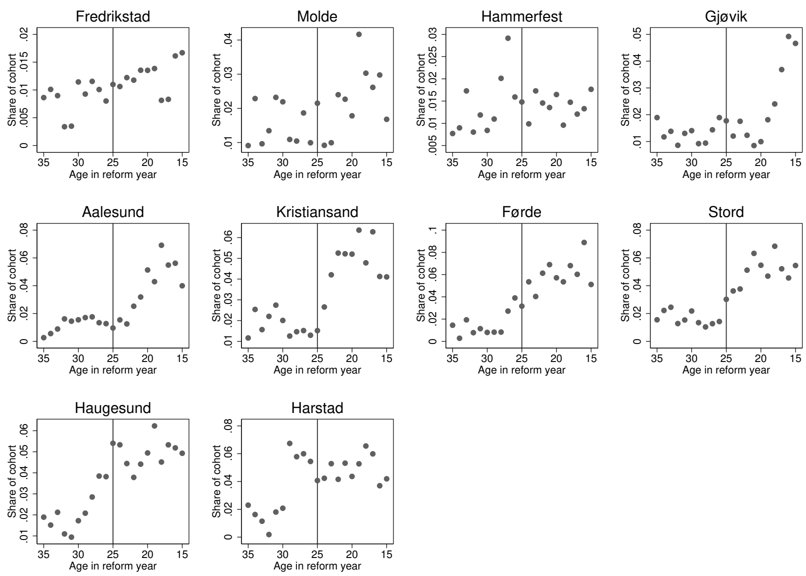


Figure B.2: Map of 46 regions, Norwegian Labour Market Regions
(Maps from Bhuller, 2009)

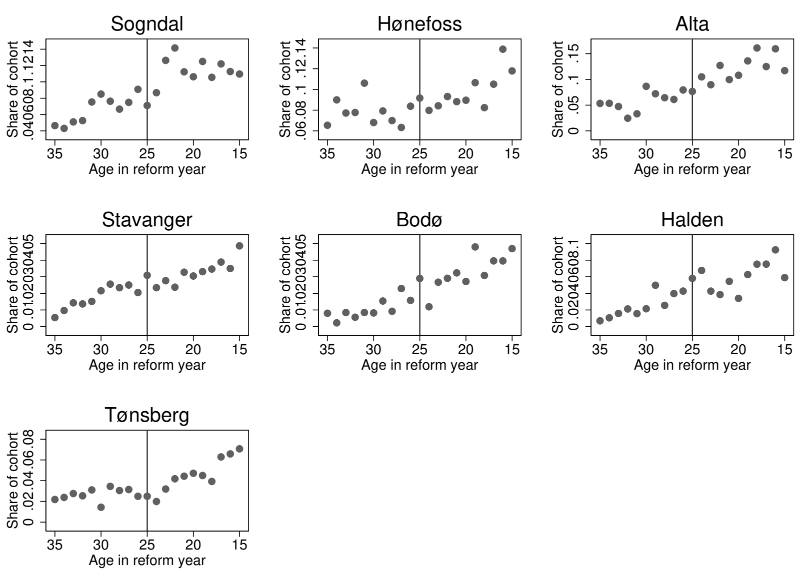
B.3 Event figures

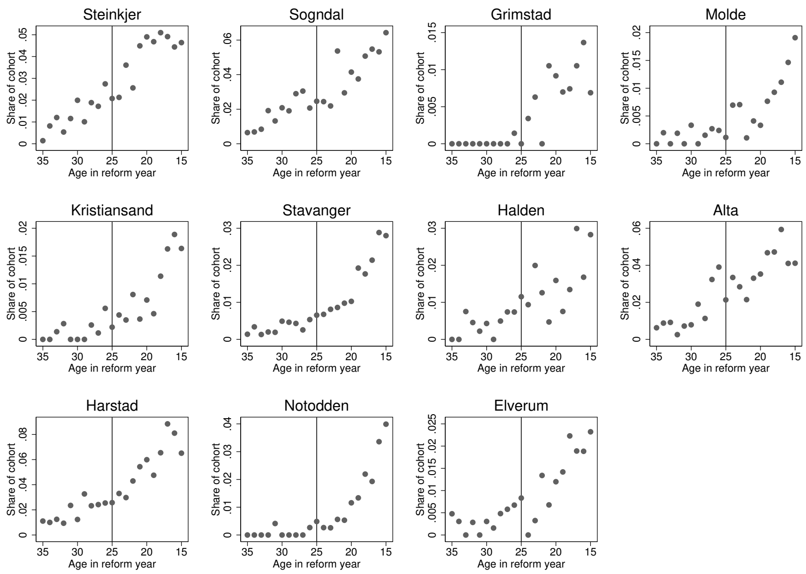
Figure B.3: The share of cohort with degree, aged 15-35 years the year of college establishment

(a) Nursing colleges

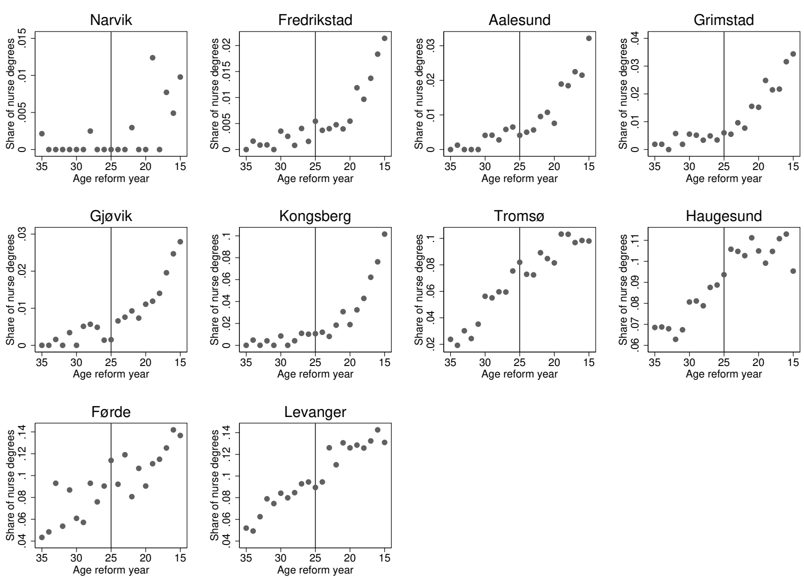


(b) Teaching colleges





(c) Regional colleges that offered degree in business



(d) Engineer colleges

Note: The figures show the share of cohort with degree in regions where a college was established. The x axis measures age in the year of college establishment while the y axis measures the share of cohort with the specfied degree.

C Regression Alternatives

C.1 Controlling for education level

Table C.1: Effect of college establishment- controlling for education level

	(i) **Nurse**	(ii) **Engineer**	(iii) **Teacher**	(iv) **Business**
College	0,0063***	0,0057***	0,012**	0,010***
(N=2309176)	(3,27)	(2,92)	(2,32)	(4,76)

Notes: t-statistic in parenthesis. Region fixed effects and cohort dummies are included in all regressions, in addition to a set of education level dummies. The standard errors are clustered at the region level. All columns present the coefficient on "Access to College" from a regression of attainment of specific college degree on access to college in home region when 25 years old, controlling for education level.
*p<0,10, **p<0,05, ***p<0,01.

C.2 Reduced form by gender

Table C.2: Second stage reduced form- Regressions of income on access to college by gender

		(i) **Nurse**	(ii) **Engineer**	(iii) **Teacher**	(iv) **Business**
Women	**College**	0,054**	-0,001	0,035	0,003
(N=1019017)		(2,33)	(-0,04)	(1,19)	(0,12)
Men	**College**	0,037*	0,001	-0,016	0,012
(N=1131923)		(1,88)	(0,58)	(-0,51)	(0,73)

Notes: t-statistic in parenthesis. Region fixed effects and cohort dummies are included in all regressions. The standard errors are clustered at the region level. All columns present the coefficient on "Access to College" from a regression of log income on access to college in home region when 25 years old.
*p<0,10, **p<0,05, ***p<0,01.

C.3 First and second stage regressions using age 19

Table C.3: First stage results: Effect of college establishment on degree attainment

		(i) Nurse	(ii) Engineer	(iii) Teacher	(iv) Business	(v) High ed.
(1)	**Baseline**	0,0076***	0,0059*	0,009*	0,0097***	0,0034
	(N=2309176)	(2,82)	(1,91)	(1,76)	(4,23)	(0,42)
(2)	**Region 46**	0,0074**	0,0061*	0,0087*	0,0075***	-0,0014
	(N=2309176)	(2,25)	(1,71)	(1,73)	(3.09)	(-0,15)
(3)	**With Controls**	0,0069***	0,0050	0,011*	0,009***	0,0012
	(N=1852402)	(2,68)	(1,53)	(1,71)	(4,26)	(0,17)
(4)	**Excl. Oslo**	0,0043*	0,0021	0,015**	0,0076***	0,0079
	(N=1746516)	(1,89)	(0,81)	(2,02)	(3,40)	(1,32)
	By gender					
(5)	**Women**	0,014***	0,0022**	0,012	0,0078***	0,0079
	(N=1124659)	(2,82)	(2,48)	(1,58)	(3,45)	(0,65)
(6)	**Men**	0,0011*	0,0092	0,0063*	0,012***	-0,0010
	(N=1184517)	(1,76)	(1,50)	(1,95)	(3,87)	(-0,16)
	By parental ed. level					
(7)	**H.E: 4 years+**	0,015*	0,004	-0,0072	0,015***	-0,0034
	(N=84563)	(1,92)	(0,56)	(0,61)	(3,43)	(-0,20)
(8)	**H.E 1-4 years**	0,0076*	0,011***	0,020**	0,016***	-0,0058
	(N=184624)	(1,72)	(2,92)	(2,31)	(3,71)	(-0,74)
(9)	**High School**	0,0069**	0,055*	0,0085*	0,008***	0,000
	(N=898258)	(2,25)	(1,71)	(1,94)	(3,47)	(0.00)
(10)	**Elementary School**	0,0045*	0,0075	0,0097	0,004***	0,0057
	(N=681267)	(1,81)	(0,17)	(1.03)	(2,97)	(0,53)

Notes : t-statistic in parenthesis. Region fixed effects and cohort dummies are included in all regressions. The standard errors are clustered at the region level. All columns present the coefficient on "Access to College" from a regression of attainment of specific college degree on access to college in home region when 19 years old. In (2) 46 regions are used, in (3) parental education and gender dummies are added, in (4) Oslo region is excluded, in (5-6) baseline regression is done by gender, and in (7-10) by parental education level.
*p<0,10, **p<0,05, ***p<0,01

Table C.4: Two Stage Least Square regressions: The effect of college degree on wage

		(i) **Nurse**	(ii) **Engineer**	(iii) **Teacher**	(iv) **Business**
(1)	**OLS**	-0,046***	0,60***	0,094***	0,37***
	(N=2150940)	(-2,78)	(92,85)	(5,72)	(40,31)
(2)	**Reduced form**	0,035*	0,0078	-0,0021	0,0016
	(N=2150940)	(1,91)	(0,34)	(-0,11)	(0,10)
(2)	**2SLS**	4,64***	1,39	-1,69	0,16
	(N=2150940)	(3,34)	(0,39)	(-0,72)	(0,10)
	F-stat first stage	8,05	3,34	3,17	17,91
(4)	**With Controls**	5,48***	1,43	-1,01	-0,19
	(N=1743826)	(4,92)	0,42	(-0,34)	(-0,09)
	F-stat first stage	7,07	2,27	2,98	18,15
(5)	**Excl. Oslo**	5,49**	-3,35	-4,80	-2,06
	(N=1675680)	(2,29)	(-0,22)	(-0,76)	(-1.07)
	F-stat first stage	3,70	2,25	4,08	11,58
	By gender				
(6)	**Women**	3,58***	7,65	-2,76	0,89
	(N=1019017)	(4,49)	(0,60)	(-0,44)	(0,32)
	F-stat first stage	8,08	6,86	2,72	11,84
(7)	**Men**	25,7*	0,56	-1,88	0,41
	(N=1131923)	(1,86)	(0,34)	(-0,44)	(0,30)
	F-stat first stage	3,07	2,19	3,68	15,04
	By parental ed.				
(8)	**H.E: 4 years+**	0,74	-4,17	-9,66	0,51
	(N=79036)	(0,59)	(-0,52)	(-0,41)	(0,37)
	F-stat first stage	3,83	0,81	0,21	11,94
(9)	**H.E 1-4 years**	6,61***	-0,50	-0,11	-0,38
	(N=174697)	(2,65)	(0,31)	(-0,31)	(-0,37)
	F-stat first stage	2,84	7,00	5,73	13,90
(10)	**High School**	5,62***	0,52	-1,10	-1,80
	(N=851781)	(4,58)	(0,21)	(-0,31)	(-0,82)
	F-stat first stage	4,97	2,68	3,77	11,96
(11)	**Elementary School**	8,73	-0,21	-2,44	-3,65
	(N=635422)	(1,82)	(-0,00)	(-0,50)	(-0,51)
	F-stat first stage	3,24	0,04	1,07	8,99

Notes: t-statistic in parenthesis. Region fixed effects and cohort dummies are included in all regressions. The standard errors are clustered at the region level. In (1) wage is regressed on having the specified degrees, (2) is the reduced form showing the coefficient from a regression of wage on access to college, (3) is the baseline IV-regression where college degree is instrumented with access to college when 19 years old, in (4) the Oslo region is excluded, in (5-6) the baseline IV- regression is done by gender and in (7-10) by parental education level.

*p<0,10, **p<0,05, ***p<0,01