CESifo WORKING PAPERS

10543
2023

June 2023

Natural Resources, Demand for Skills, and Schooling Choices

Aline Bütikofer, Antonio Dalla-Zuanna, Kjell G. Salvanes

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 10543*

Natural Resources, Demand for Skills, and Schooling Choices

Abstract

This paper studies the consequences of the buildup of a new economic sector—the Norwegian petroleum industry—on investment in human capital. We assess both short-term and long-term effects for a broad set of educational margins, by comparing individuals in regions exposed to the new sector with individuals in unexposed regions. Importantly, we analyze how the effects and the mechanisms change as the sector develops. Our results indicate that an initial increase in the high school dropout rate is short-lived both because dropouts get their degrees later as adults, and because later-born cohorts adapt to the new needs of the industry by enrolling more in vocational secondary education. We also observe a decrease in academic high school and college enrollment except for engineering degrees. Financial incentives to both completing high school and field of study, are the most likely channels driving these effects.

JEL-Codes: J240, J230, I260, I230.

Keywords: school choice, demand for skills, natural resource sector.

Aline Bütikofer
Department of Economics
Norwegian School of Economics
Bergen / Norway
Aline.Buetikofer@nhh.no

Antonio Dalla-Zuanna
Bank of Italy / Rome / Italy
antonio.dallazuanna@bancaditalia.it

Kjell G. Salvanes
Department of Economics
Norwegian School of Economics
Bergen / Norway
Kjell.Salvanes@nhh.no

June 30, 2023
The authors gratefully acknowledge comments by Sandra E. Black, Jeffrey Wooldridge, and seminar participants at the Institute of Fiscal Studies, Paris School of Economics, the Norwegian School of Economics, the Bank of Italy, the 2019 IWAEE Conference, and the 33rd EALE Conference (Padua). This work was partially supported by the Research Council of Norway through project No. 316085 and through its Centres of Excellence Scheme, FAIR project No. 262675, and by the NORFACE DIAL grant 462.16.159.

1 Introduction

The development of new sectors due to technological shocks, trade, or the discovery of natural resource reserves leads to major changes in the labor market. In turn, these changes may affect incentives for investment in human capital—a crucial factor for future productivity, growth, and welfare, as well as the ability to adjust to future transformations of the economy. Such economic restructuring can both be temporary, such as resource price booms, or longer-term, such as trade relations and technological change or the development of large, resource-based industries.

So far, the literature has largely analyzed the short-run effects of economic shocks on schooling. On the one hand, examples are changes over the business cycle which include temporary shocks induced by resource price changes (see, e.g., Black et al., 2005a; Emery et al., 2012; Morissette et al., 2015) or housing booms (Charles et al., 2018) that allow identifying effects during relatively short boom and bust cycles. On the other hand, the literature has focused on the analyses of short-run effects of the establishment of new industries, such as the fracking industry (Cascio and Narayan, 2022) or the establishment of export industries (Atkin, 2016). Hence, these studies do not provide a full understanding of the consequences of the development of a new economic sector that leads to permanent labor market changes, or the dynamics of human capital investments over time. Importantly, establishing new industrial sectors may undergo different phases as new technologies need to be developed before the sector matures and becomes an integral part of the economy. Thus, the pattern of human capital investment may also change. Due to the introduction of new technologies as industries mature, the demand for skills changes, thereby affecting different educational margins. While low-skilled workers have been the main focus of the literature studying short-run effects of economic shocks, long-term developments may, for example, alter the incentives to return to school or to choose a specific field of study at the university level. This longer-run and more complex role in the development of new economic sectors is not well understood. Nevertheless, the composition of education in a labor market is essential to adjustments to future economic shocks and it is therefore important to understand also the longer-term consequences of long-lasting economic changes on human capital accumulation.

In this paper, we investigate the effect of oil discoveries in Norway in the late 1960s and the resulting buildup of the Norwegian oil sector on educational choices. We study different stages of the Norwegian oil sector starting before the first oil discovery in the 1960s, through periods of rapid technological changes in the 1980s until today. Moreover, we focus on different educational choices including high school completion, vocational vs. academic high school, and field of study at the university level. Four features of our setting make it ideal for analyzing the long-term dynamic impact of a large sectoral restructuring of an economy on human capital investments across different educational margins. First, the activity is concentrated in the labor markets in the areas close to the newly discovered oilfields, allowing us to identify the timing and the geographical dimension of the change by separating affected and non-affected areas and cohorts (see, e.g., Feyrer et al., 2017; Bütikofer et al., 2022). Second, the restructuring started in the late 1960s allowing us to have a long enough period to analyze the long-term development of the effects and the mechanisms. Third, oil extraction in Norway is technologically challenging as it takes place offshore and underwent three main phases. In the 'early phase' (1970-1975), the core technologies for extraction were mostly imported from abroad. This import-heavy first phase was followed by an 'expansion phase' (1976-1985) when new and advanced technologies were developed locally for deep sea drilling, and finally, by a 'mature phase' after 1986, when the Norwegian oil sector became the dominant sector in Norway. These three distinct development phases suggest that the effect of the oil discovery may not be uniform over time but affects low-skilled workers more in the early phases and high-skilled workers or higher education decisions in later phases. Fourth, we use a panel of individual-level population-wide register data for cohorts born between 1946 and 1984 that allow us to identify different school choices (and earnings), including dropping out and reentering high school, as well as fields of higher education. We specify a dynamic difference-in-differences framework using cohorts who make educational decisions before the first oil discovery in Norway as the baseline. Importantly, we also analyze the mechanisms driving the educational choices over the development of the oil sector and across educational margins.

Conceptually, our analysis is guided by the notion that, for an individual, each schooling path has an expected return, but may also entail an opportunity cost if attending school prevents the individual from earning a salary (Mincer, 1958; Becker, 1964). Students compare the returns and costs of the alternatives they face and then select the education type with the highest expected net return. As discussed in Atkin (2016), changes in the local industrial structure may modify both returns and costs, potentially giving opposite incentives. Hence, it is hard to make clear predictions about the impact of an economic shock on educational attainment. Since the opportunity costs of schooling are immediately revealed, while the expectation of the returns to education requires several years of observation to reveal itself, there may be some misalignment in the timing of industrial restructuring and its effect on the two opposing forces. A dynamic approach is therefore crucial to understand the mechanisms at play. Moreover, our approach is relevant if we are to understand whether individuals re-adjust their education if, in the long run, the differences between these two forces evolve. In addition, the expected returns to educational investments may also evolve as the oil sector underwent different phases. For this reason, when we investigate the dynamics of educational investments, we do so at different educational margins and not just focus on high school dropouts (Cascio and Narayan, 2022) or the major choice at college (Weinstein, 2020).

In addition to assessing the dynamics in the long run, we extend the literature on the effect of resource shocks on education along two dimensions. First, the impact may not just be on the level and length of education students decide to take, but also on the field of specialization. Hence, limiting the analysis to the quantity of education may miss some aspects of the effect of a large change in the labor market on human capital accumulation, whereas looking at the content or quality of education may reveal a more nuanced picture (see, e.g., Weinstein, 2020). In particular, in the context of an economic restructuring that includes technological change, the choice of field of study may not only be consequential on the tertiary educational level but also the secondary education level when students decide to enroll in academic or vocational high school tracks.[1] Second, in the medium and long run, sector-specific shocks may spill over to other sectors in the local labor market (Black et al., 2005b), changing educational choices for a much broader range of individuals. Hence, understanding the overall impact of large economic changes on the educational decisions of individuals requires a longer-term perspective that also captures changes in incentives for non-directly affected individuals.

We report six core findings. First, we show that in the immediate aftermath of the first oil discovery, the proportion of male students dropping out of high school increases significantly in the areas most affected by the oil discovery. This finding is largely consistent with findings from the US and Canada (see, e.g., Cascio and Narayan, 2022; Kovalenko, 2023; Morissette et al., 2015). Studying the underlying mechanisms, we show that this is due to an increase in the opportunity cost of attending high school. However, the regional differences in high school dropout rates even out for cohorts making educational decisions after the initial phase. Second, we document that the increase in the high school dropout rate is a temporary phenomenon for the cohorts that initially quit high school due to the oil discovery. Studying educational outcomes ten years after these cohorts started high school documents that individuals completed high school with some delay and that there is no permanent effect of the oil discovery on the high school completion rate. This finding is consistent with the results of Emery et al. (2012). Third, throughout all phases, we document a large increase in the number of male students in high-oil regions attending vocational high school and a decrease in students attending academic high school. Analyzing the mechanisms, we show that earnings for individuals who graduate from a vocational high school increase substantially throughout the period relative to individuals with academic education in high-oil regions. Fourth, the observed decrease in the share of men attending academic high school has long-term consequences: the proportion of university graduates in high-oil regions decreases permanently. Fifth, the decrease in university graduates is not uniform across fields of study. Some fields of higher education related to the oil sector, such as engineering, are affected positively by the oil boom. Changes in the return to educational degrees are also here the likely mechanism. While the returns to completing a university degree relative to obtaining a vocational high school degree decreased, the returns to graduating in engineering as compared to other university subjects increased. The returns to other fields of study, such as business or law degrees, are not affected. Finally, we document that women adapt differently than men to the buildup of the oil sector. While the educational choices of women are not directly affected by the initial oil discovery, women are also more likely to enroll in vocational high school education as the industry matures likely due to the spillover effects of the oil boom on the service sector in high-oil regions. To understand the mechanisms behind the six main results, we also estimate the elasticities of changes in educational attainment to changes in earnings and returns for different education types and compare these values to the estimates obtained in the past for short-lived sector-specific booms (Black et al., 2005a; Morissette et al., 2015; Charles et al., 2018).

[1] Note that the Norwegian School system differentiates between vocational and academic secondary education, as is common across most OECD countries (OECD, 2019).

While high school dropouts seem to respond similarly to longer and short-run economic shocks, educational choices at the university level show a stronger reaction to more permanent changes.

The paper is structured as follows. In Section 2, we describe the oil sector buildup and its long-term evolution and provide insight into potential channels through which the new sector's buildup may have impacted Norwegian students' educational choices. Section 3 details our empirical strategy and the data used. In Section 4, we present our main results and provide a battery of robustness tests. Section 5 investigates the mechanisms and Section 6 analyzes the presence of spillover effects from the oil sector to other sectors. Section 7 concludes.

2 The Norwegian Oil Industry and Education System

2.1 The Norwegian Oil Sector

The oil and gas industry is currently Norway's most important industry in terms of both treasury revenue and investment (Ekeland, 2015). However, the first oil discovery on the Norwegian continental shelf was only made in 1969 when Ekofisk, one of the largest offshore oil fields ever found, was discovered. The offshore exploration activity was dominated by foreign firms, with little spillover effects on the mainland labor market.

We identify three distinct development phases after the first oil discovery: the early phase (1970-1975), the expansion phase (1976-1985), and the mature phase (1986-). During the early phase, core technologies for extraction were mostly imported from abroad. Norwegian involvement in the sector came through already existing shipping and engineering companies, which converted to supplying the oil sector, especially in the areas close to the newly discovered offshore oilfield (in south-western Norway around the city of Stavanger). The expansion phase is characterized by requests for a more protectionist approach. Hence, the Norwegian government played an active part in developing local expertise and skills. Statoil, a state-owned company, was central to this process. The company contributed by concentrating oil activities around its headquarters in Stavanger. Moreover, the government decided that all oil from the Norwegian continental shelf needs to be landed on the Norwegian shore which led to the buildup of supply bases and refineries. Finally, during the mature phase, the protectionist policies came to an end as Norwegian companies could compete internationally in terms of technological know-how and skills.[2]

To provide some background to the reallocation of labor following the breakthrough of the oil industry, Figure 1 shows the proportion of men aged between 30 and 35 years employed in the primary, secondary, and tertiary sectors in 1960, 1970, 1980, 1990, and 2000. The tertiary sector expands rapidly during the second half of the twentieth century almost doubling the fraction of young men employed in this sector (Panel (a)).[3] Most of the industries involved in the oil sector

[2] A more detailed description of the evolution of the Norwegian oil sector and the three phases is provided in Online Appendix A.

[3] Figure A1 displays the same numbers for women. The fraction of young women employed in the tertiary sector also increased (Panel (a)), despite already being high (70%) in 1960.

are concentrated in the primary (extraction activity) and secondary sectors (e.g., manufacturing of oil products, machinery, and ships). The oil sector increased from about 7.8 percent in 1960 to 13 percent in 2000 (Panel (a)). Note that our definition of the oil sector includes both extraction activities and supply industries.[4] Hence, the definition is sufficiently broad that it is no zero in the years preceding the first oil discoveries. Female employment in the oil sector remained low. Around 1990, the share of women was about half that of men. These gender differences have been noted previously in other countries, such as the US and Canada (Cascio and Narayan, 2022; Emery et al., 2012). For this reason, we mainly focus on men in the remainder of the paper. An exception is Section 6 where we focus on women when analyzing spillover effects from the oil sector to other sectors.

There is clear evidence of a geographical concentration of oil activities in some specific local labor markets (LLMs, see Figure A2).[5] We, therefore, define 'low-oil' regions as labor markets where less than 7.5% of the total employment was in the oil sector in 1980 and 'high-oil' regions as labor markets where more than 10% of total employment was in the oil sector.[6] These geographic differences are partly due to the exogenous distance between the labor markets and the newly discovered oilfield (see, e.g., Cust and Harding, 2020), partly to the region's pre-oil industrial composition, and partly for political reasons (e.g., the establishment of Statoil's headquarters in Stavanger). Panel (b) of Figure 1 shows that high-oil regions were less service-oriented already in 1960 and 1970. Nevertheless, the difference in employment in oil-related industries only explained a very small fraction of the difference in non-service employment at the start of the oil sector buildup (Panel (c) of Figure 1). Since 1990, however, the oil sector mainly explains the (growing) difference in non-service employment between the two regions.

2.2 Returns to Education, Opportunity Costs, and the Norwegian Oil Sector

Educational decisions may be affected by a combination of the oil discovery-induced changes in returns to education and opportunity costs of studying (Becker, 1964; Mincer, 1984). The standard models for labor demand and supply predict that there will be an increase in wages for workers with skills that are in demand in the new sector. This, in turn, may lead to an increase in the enrolment rate for education programs offering these skills, as forward-looking students consider the potential returns on their educational investment. We refer to this mechanism as the returns to education channel. However, this is not the only mechanism in play. If, as a consequence of the oil discovery, wages increase for low-level education (e.g., high school), moving to the next educational level (e.g., university) costs more in terms of foregone earnings. We refer to this mechanism as

[4]Using three-digit industry codes, the oil industry includes crude petroleum and natural gas production, petroleum refining, manufacturing of petroleum and coal products, manufacturing of machinery, which includes the manufacturing of oil and gas well machinery and tools, the manufacturing of transport equipment, which includes the building of ships and boats, and other construction, which includes oil well drilling. See Bütikofer et al. (2022); Brunstad and Dyrstad (1997) for a more detailed discussion of the definition of the Norwegian oil sector.

[5]Following Bhuller (2009) we can identify 46 labor markets, which span the whole Norwegian territory and are an aggregation of municipalities (lowest administrative level) defined based on commuting patterns.

[6]7.5% is the median oil employment level in 1980 across labor markets, while 10% is the top quartile oil employment. The average oil employment share across low-oil regions is 5.5%, while it is 13% for high-oil regions.

the opportunity cost channel. The opportunity cost channel may offset the potential positive effect on enrolment of an increase in returns to high-level education. At the same time, the opposite may be true. Even if earnings for low-educated workers increase, this may not lead to an increase in the share of individuals with low-level education if the shift is accompanied by changes in the returns to higher-level education.

2.2.1 The Norwegian Education System

In this section, we briefly describe the Norwegian education system and the crucial junctures when students are required to decide their educational path. We then provide some descriptive evidence of the difference in education between workers in the oil and non-oil sectors, to give an impression of the types of education that are likely to be in higher demand, following the expansion of the oil sector.

The Norwegian education system consists of four levels: primary school, middle school, high school, and university education. Compulsory school ends with middle school at age 16. High school lasts three years and the system comprises a vocational and an academic track. Academic high schools are geared toward future enrolment in a university. Vocational high schools lead to employment in a given occupation (e.g., electrician) and are based on a combination of school and apprenticeship. With its two tracks, the Norwegian high school system is very similar to most other European countries and Australia. Every student is allowed to enroll in a high school track of her choice. The rules for assigning students to each specific high school are decided at the county level and are usually based on geography and/or performance. Once students obtain an academic high school diploma, they may decide to enroll in a specific field of study at university. The decision to enroll at university is thus made at the age of 19 years for students graduating high school in the normal time. Admission to a combination of a specified field of study and institution (e.g., law at the University of Oslo) is based on high school GPA (a combination of teachers' grades and centralized exit tests), and admission to degree programs in high demand, such as medicine and the science, is very competitive.[7]

Based on the structure of the Norwegian schooling system, we identify four crucial educational margins in our setting. The first is at age 15 when students have to decide what type of high school track to enroll in. The second is high school completion where we distinguish between completing high school before age 23 or later as an adult to understand whether the economic restructuring changed the timing of high school completion. Third, we analyze whether the oil sector buildup affected enrolment in higher education. This decision is taken at the age of 19 years for most students. Finally, the labor market conditions have been shown to affect the field of study students select at the university in the US (Weinstein, 2020). Hence, we also investigate this additional margin, focusing on fields of study relevant to the oil industry.

[7] Note that primary schools are administered at the municipality level, high schools at the county level (above the level of municipalities), and college education at the national level. The financing of all education is decided at the national level. Most students enroll in public institutions on all educational levels, and the few private institutions that do exist are also largely funded and regulated by the state. Moreover, the high school education system for adults is regulated and financed at the national level since the 1970s (Bjerkaker, 2016).

2.2.2 The Oil Sector and Education

Norway experienced a sharp increase in the percentage of individuals enrolling in secondary and tertiary education since the 1970s (Figure A3). This trend is similar to trends in most Western countries and driven by individuals obtaining an academic high school degree.

In Panel (d) of Figure 1, we compare the educational attainment of young workers employed in the oil sector and the full population. The fraction with an academic education has remained smaller in the oil sector, highlighting the fact that the oil sector offers a substantial amount of low-skilled jobs. As the number of students who complete high school increased in the 1980s and 1990s, sectoral differences in the share of employees without high school degrees diminished, and differences in the proportion of students completing vocational education emerged. This descriptive evidence supports the hypothesis that there continued to be a higher demand for non-academically educated workers following the expansion of the new industry (see, e.g., Black et al., 2005a; Cascio and Narayan, 2022). Moreover, Figure 1 shows a decrease in the gap in the proportion of academically educated workers until 1990 between the oil sector and the whole economy. This convergence may be due to an overall increase in the supply of university graduates or driven by the new technologies developed for deep-sea drilling during the mature phases of the oil sector that increased demand for high-skilled workers such as engineers.

To summarize, our descriptive evidence shows that the establishment of the oil sector in Norway underwent different phases, with a heterogeneous geographical impact. In the next section, we describe our empirical strategy to causally identify the impact of the oil sector on educational choices.

3 Empirical Strategy and Data

Our empirical strategy to isolate the impact of the new industry on individuals' human capital accumulation exploits both the geographical variation in the importance of the oil sector across labor markets and the timing of the different development phases of the oil industry. This enables us to build a dynamic difference-in-difference framework and estimate both the short- and the long-run effect of the new industry, while also investigating pre-trends in outcomes.

3.1 Treated and Control Areas

In our main specification, we compare the high-oil labor markets described in Section 2.1, i.e., labor markets where, in 1980, more than 10% of the labor force was employed in the oil sector, to the low-oil labor markets, where employment in the oil sector was less than 7.5%. We prefer this comparison to a continuous measure of oil impact because the distribution of oil employment across labor markets is non-linear, especially at the top (see Figure A4). In Section 4.1.2, we perform several tests to confirm that our results do not depend on the definition of oil regions we have used, and we also test for continuous measures. To balance accuracy in terms of capturing the circumstances that affected schooling decisions and avoiding endogeneity in the selection of migration into regions driven by the oil sector, we assign individuals to their oil region of birth. We test the robustness of our results to alternative assignments in Section 4.1.1, where we also show that migration toward high-oil areas is low even among the most recent cohorts.[8]

This strategy relies on the assumption that the evolution of the educational choices of students born in high-oil areas would have followed the same path as is observed in the low-oil areas in the absence of the oil discovery. Below, we empirically demonstrate that no significant differences existed between regions in the dynamics of educational choices before the first oil discovery. However, we can envisage reasons why students in areas that are not directly affected by the oil discovery might also experience some spillover impacting their educational choices. In particular, while oil jobs were concentrated geographically in the first years following the discovery, it is possible that, in the longer run, more oil-related jobs were created elsewhere in the country, because of the changes in the nature of the oil jobs or because of new oil discoveries along the coast. In the data, we observe that the proportion of workers employed in the oil sector in low-oil areas remained small at least until 2003 when the youngest cohort we consider was 19 years (oil employment was 5.5% in low-oil regions and 15.6% in high-oil regions). All in all, the direct effect of the oil industry seems to be substantially smaller in low-oil regions throughout the period. However, if there is any effect, our results should be interpreted as a lower bound.

3.2 Affected Cohorts

Our data enable us to investigate the completion of education among individuals in high- and low-oil areas across different cohorts over the development of the oil sector. In addition, we can observe cohorts that made their educational choices before the first oil discovery. To identify which cohorts are affected by the oil sector, we specify at what age an individual makes her educational choices. Based on this age, we assign individuals to one of the three development phases of the oil sector. For example, cohorts born in 1960 are assigned to the early phase concerning their decision about whether to enroll in a vocational or academic high school (they were 15 in 1975) and to the expansion phase for the university decision (they were 19 in 1979). Likewise, individuals born in 1955 were in their late teens when oil was discovered and made their decision about the type of high school before the oil discovery, while they could adjust their field of study at the university to the new labor market composition. Hence, they are assigned to the early phase concerning tertiary education decisions. Given that the first cohort we observe was born in 1946 (24 in 1970), we have pre-oil cohorts for all the education margins we study and we exploit these observations to analyze pre-trends across the different regions.

3.3 Empirical Specifications

In our main specification, we compare the educational choices of different cohorts born in high- and low-oil areas, utilizing a dynamic difference-in-differences framework where we control for

[8]This strategy follows Bütikofer et al. (2022), who perform several tests to assess the comparability of the different areas before the oil discovery in terms of income per capita, earnings dispersion, intergenerational persistence of earnings and education, and returns to education, and show that there was no difference in any of these factors before the first oil discovery. Since the focus of this paper is on the educational choices of individuals, unlike Bütikofer et al. (2022), we choose an empirical specification that allows testing whether the trends in education in the treated (high-oil) and control (low-oil) areas were parallel before the discovery of the first oil field.

labor market and cohort fixed effects. We carry out the analysis at the individual level. Hence, we estimate a linear probability model for each educational outcome, which corresponds to all the margins of adjustment in terms of the educational choices of individuals:

$$edu_{i,c,llm} = \alpha + \sum_{c=1946}^{1984} \theta_c Cohort_{c,i} \times HighOil_i + \gamma_c Cohort_{c,i} + \gamma_{llm} LLM_i + \beta x_i + u_{ic}. \quad (1)$$

$edu_{i,c,llm}$ is a dummy equal to 1 if individual i born in year c in the local labor market denoted llm, obtained education edu. $HighOil_i$ is a dummy for being born in the high-oil region. $Cohort_{c,i}$ and LLM_i are cohort and labor market fixed effects. x_i are individual characteristics, including the population in the municipality of birth in the year the person turns 15, the number of siblings, the birth order, the annual earnings of fathers in the year the individual turns 17, and the highest educational attainment of parents (controlling for family-level characteristics may be particularly relevant for recent cohorts, whose parents were also affected by the oil sector during their educational career). To increase precision, we combine cohorts in groups of two (for example, cohorts born in 1946 and 1947). We cluster the standard errors at the labor market level.[9]

Coefficient θ_c is the estimate, for each cohort, of the difference in the proportion obtaining education level edu between the high- and low-oil regions. We exclude the dummy for the interaction between $HighOil$ and $Cohort_{1948}$, so that the difference between areas for cohorts born in 1948 and 1949 is the baseline. These cohorts were already in their 20s when oil was first discovered and they were thus finished with their core educational choices. When reporting the results, we highlight for each cohort the phase of the oil sector buildup during which they took their educational decision, to document the link between the regional differences in education and the development phase of the oil industry. As mentioned above, because educational choices are taken at different ages, the same cohort may be assigned to different phases depending on which education margin we are considering.

To further quantify the differences during each phase, we also estimate Equation (1), assigning each cohort to a dummy for the development phase in which they make their decision about each education level. We thus interact the $HighOil$ dummy with three dummies, one for each development phase, controlling for the phases' main effects. We thereby estimate the following regression:

[9] Because we assign treatment based on the labor market of birth and we exclude some labor markets—those with oil employment between 7.5 and 10% in 1980—our empirical strategy can be interpreted as an experiment where treatment is assigned based on the labor market. For this reason, the clustering at the labor market level is justified (Abadie et al., 2022). In the regression, we include 40 labor markets. This clustering takes into account the fact that the standard errors are likely to be correlated within labor markets in the cross-section and also serially correlated (Bertrand et al., 2004). However, since our time dimension is large, serial correlation "dies out" when periods are far apart, something we do not impose when using cluster-robust standard errors at the labor market level (Wooldridge, 2003). In Table A1, we present the estimates of an alternative specification of Equation (1), where observations are collapsed at the labor market level, parameters are estimated using OLS weighting for population size, and we apply Newey-West standard errors, assuming autocorrelation of order 1 in the standard errors within labor markets.

$$
\begin{aligned}
edu_{i,llm} =& \alpha + \theta_{Early} Early_i \times HighOil_i + \theta_{Expansion} Expansion_i \times HighOil_i \\
&+ \theta_{Mature} Mature_i \times HighOil_i + Early_i + Expansion_i + Mature_i \\
&+ \gamma_{\mathbf{llm}} \mathbf{LLM_i} + \beta \mathbf{x_i} + u_{ic}.
\end{aligned}
\tag{2}
$$

Parameters θ_{Early}, $\theta_{Expansion}$ and θ_{Mature} are the difference between high- and low-oil regions in the proportion of students obtaining education level *edu* for the cohorts deciding about *edu* during each of the three development phases, compared to the baseline, pre-oil discovery difference.[10]

3.4 Data

Norwegian administrative data from different administrative units, such as tax authorities and educational institutions, are our primary data sources. We also use centennial census data starting in 1960. We exploit several features of these data. First, the long panel structure enables us to identify where and when people are born and to follow them for several decades in the data. In addition, we observe medium- and long-term educational outcomes on an individual level. Second, the education register contains details on the start date, end date, and which fields are chosen in high school and higher education. Third, we are able to match individuals to firms and sectors as working adults (from their first job). Last, since we have a population register including parental identifiers, we can link parents and children and establish an individual's socioeconomic background.

We focus on individuals born between 1946 and 1984. The size of the sample used in the analysis of each educational outcome in different phases and regions is reported in Table A2. Table A3 shows the proportion of individuals completing the different education levels by development phases and oil regions. This table confirms the general expansion of secondary education and, in particular, of academic education throughout the analysis period. A detailed description of the different variables and data is in Online Appendix B.

4 Results

This paper investigates the impact of the Norwegian oil sector buildup on six different education margins: (i) high school dropout by age 23, (ii) high school dropout by age 30, (iii) vocational high school completion, (iv) academic high school completion, (v) university degree, and (vi) attainment of an engineering degree when at university. We exploit the empirical specification described in Section 3.3 to estimate the magnitude of the effect of the new sector on educational choices. Figure 2 plots the estimates and confidence intervals for coefficients θ_c from Equation (1) for all different educational outcomes, i.e., the cohort differences in the proportion obtaining

[10] As in Equation (1), we cluster the standard errors at the labor markets level to take into account correlations within labor markets. To avoid problems of serial correlation in the error term, we also estimate the equation, collapsing the observations at labor market times phase cell (160 cells), weighting each cell by the respective population, and computing heteroskedasticity-robust standard errors (Bertrand et al., 2004; Emery et al., 2012). The results of this exercise are shown in Table A5.

each education between high- and low-oil regions as compared to the pre-oil cohorts born in 1948 and 1949. The dashed vertical lines separate the development phases during which the different cohorts make different educational choices, taking into account the fact that the same cohort takes different decisions at different points in time (and thus is assigned to different phases depending on the educational choice we are examining). Estimated parameters and standard errors for the estimates of Equation (1) are also reported in Table A4. Alternatively, in Table 1, we instead show the estimated parameters of the regression using three phase-dummies and directly assigning each cohort to a development phase. Importantly, the estimates for all the outcomes confirm that there was no differential trend between regions for cohorts making educational choices before the first oil discovery, thus supporting the parallel trend assumption and the causal interpretation of our results.

High School Dropout Rate and Completion with Delay. Panel (a) of Figure 2 shows a statistically significant increase in the high school dropout rate (not split by track, measured at age 23) among students who were between 15 and 20 years in the early phases of the oil boom. The increase in the high school dropout rate at the onset of a resource boom is in line with findings on price-induced natural resource booms or the buildup of the US fracking industry (see, e.g., Black et al., 2005a; Cascio and Narayan, 2022). Averaging across early-phase cohorts, this is an increase of 1.8 percentage points (Column (1) of Table 1), which, compared to the low-oil areas, corresponds to a 3.5% increase in the high school dropout rate. However, this increase in the dropout rate is a short-lived effect. Later-born cohorts in the high-oil areas experience a small and insignificant decrease in the high school dropout rate relative to individuals in the low-oil regions.

Nevertheless, the increase in the dropout rate for the cohorts in high school right after the first oil discovery may be temporary. As the oil sector evolved, the incentives to complete high school as an adult might change. Panel (b) of Figure 2 and Column (2) of Table 1 show the effect on the high school dropout rate measured at age 30. The increased dropout rate in high-oil areas which we observed at age 23 vanishes by age 30. This implies that early-phase high school dropouts indeed completed high school with some delay.[11] This result is in line with Emery et al. (2012), which showed that men postponed completion of post-secondary degrees during the Canadian oil boom, with no long-run effect. When grouping all cohorts in the expansion phase, we also estimate a marginally statistically significant decline in the proportion of dropouts by age 30 in the high-oil regions, which indicates that completing high school may have increased in the more recent development phases, thanks to students going back to school between age 23 and 30.

High School Track. In Panels (c) and (d) of Figure 2, we investigate the difference in enrolment in vocational and academic high schools. Looking at high school completion by age 30, when little differences in the dropout rate are observed, there is a large and significant increase in vocational school completion in high-oil regions, starting from the expansion phase. Importantly,

[11] Exploiting the SUR framework, we tested for equality in the parameters for the interaction term between the high-oil dummy and the dummy for the early phase in the regression for completing high school by age 23 and by age 30. The formal test supports the two parameters being different at the 5% level.

in this phase of the oil industry, academic high school completion, on the other hand, declined (see Column 4 of Table 1). The increase of 5.2 percentage points in vocational attainment during the expansion phase, and of 3.9 percentage points during the mature phase, corresponds to an increase of between 30 and 25%. Academic high school completion drops by 2–3 percentage points, which corresponds to a decline of approximately 4% compared to the baseline. Overall, also considering the decline in the high school dropout rate observed during the expansion phase, the steep increase in vocational education seems to draw from two educational choices: students who would otherwise not have completed high school and students who would have taken an academic track.

University Attainment and Choice of Major. Panel (e) of Figure 2 shows that there are no significant regional differences in university completion in the early phase, only a slight reduction for cohorts making their choices towards the end of the expansion phase, which then extends into the mature phase. When we group all cohorts belonging to the same phase, the decline in university enrolment is only statistically significant for students in the mature phase. Column (5) of Table 1 documents a decline of 2.6 percentage points or 6% in university enrollment. This documents a relatively late response, especially compared to previous research on the effect of short-lived resource booms (Emery et al., 2012). Comparing Panel (e) with Panels (c) and (d) of Figure 2, we observe that the decline in university attainment starts with cohorts that also have less academic high school attainment. This evidence supports the hypothesis that the decline in university graduates has its roots in the educational choices that students make before starting high school.

On the other hand, we also find that the proportion of university graduates who chose engineering as their field of study increased quite dramatically for the same cohorts, by 4.6 percentage points on average, or 40% (Panel (f) of Figure 2 and Column (6) of Table 1). Interestingly, the increase is also present at the population level (Panel (b) of Figure A5), even though the overall university graduation rate decreased slightly during this period.[12]

4.1 Robustness Tests for Treatment Definition

We perform several robustness checks to establish whether our results depend on how we define individuals who are affected by the new oil sector. We first show that changing how we assign individuals to different oil regions has little impact on the results, mostly because of the limited geographical mobility that followed the oil discovery. Second, we change the definition of the geographical incidence of the oil sector to show that our findings do not depend on the thresholds we impose to identify high- and low-oil regions.

4.1.1 Assignment of Local Labor Markets

In our main analysis, we assign individuals to labor markets and, consequently, the oil area based on their municipality of birth to avoid capturing the effect of educational choices of endogenously

[12] For the sake of completeness, we investigated whether other subjects such as law or business display similar increases as engineering degrees. We found that, if anything, both fields decrease their enrolment rate in high-oil areas following the oil discovery, in line with the general decrease in university enrolment.

selected migrants or their children. However, in the presence of substantial migration, the birthplace may not be the relevant labor market for the student. In addition, individuals born after 1970 may well be the children of workers who decided to migrate to the oil areas as a consequence of the oil discovery. In this section, we first show that internal migration was very low, then re-run our core analysis, assigning individuals to labor markets based on place of residence at the age of 15 and the mother's birthplace, and show that the results are largely unchanged.

Before the first oil discovery, there was an increasing trend of people moving from the high-oil to low-oil regions (a large share explained by people moving to Oslo, Figure A6). After the first oil discovery, migration to the high-oil regions became more common, while the proportion of individuals moving in the opposite direction remained almost constant. Only 2 to 4% of the children born in the high-oil region after the first oil discovery had parents born in the low-oil area. This share was lower than the fraction of parents migrating in the opposite direction for most of the cohorts we study.

Figure 3 plots the point estimates of parameters in Equation (1) when assigning students to the place where they live when they are 15 years and to the place of birth of the mother. We choose age 15 as it is the most relevant in terms of high school education decisions and until 15 years individuals still live and move with their parents. The mother's place of birth, on the other hand, should not be affected by the oil discovery or her later moving decisions. Table 2 reports the corresponding estimates of the regression using the three phase-dummies. The results are largely unchanged for cohorts born after 1970. The small differences in estimates are in line with the evidence presented by Bütikofer et al. (2022), who explain the low geographic mobility by redistribution policies, the increase in house prices in high-oil regions, and the positive moving costs for people who would have gained most by relocating to the high-oil regions (see also, e.g., Bartik, 2017).

4.1.2 Oil Region Definition

Next, we investigate how robust our results are to changes in the geographical definition of the impact of the oil sector. First, we estimate the regression when we group cohorts in the three development phases, interacting these dummies with a continuous measure of the employment rate in the oil sector in 1980 instead of more arbitrary threshold values. Column (4) of Table 2 shows that the signs of the estimates are the same and the magnitudes are also comparable. For example, this model predicts a difference between the Stavanger labor market (high-oil, 11.6% oil employment in 1980) and the Oslo labor market (low-oil, oil employment 4.8%) in vocational high school enrolment during the expansion phase of approximately 4.6 percentage points, not far from the 5.4 percentage point difference between high- and low-oil regions estimated in our main specification.

In addition, in Figure 3 (and in the last two columns of Table 2), we show that estimates of the parameters of Equation (1) do not depend on how we define the threshold for low- and high-oil areas. In particular, we report the estimated differences between high- and low-oil areas when we change the threshold values for the definition of the oil sector; instead of 7.5 and 10% of 1980 oil employment, we use 6.5 and 9% (down) and 8.5 and 11% (up). None of these changes impacts our results substantively.

5 Mechanisms

Our findings support the oil sector buildup has had quite substantial effects on educational decisions along many margins and that the effects evolve over time as the industry develops. In particular, we report five main changes in educational choices: (i) a sudden increase in the high school dropout rate that does not persist in the long run, (ii) an increase in vocational high school graduation, with (iii) a slight decrease in academic high school graduation, which led to (iv) a decrease in university attainment, and (v) an increase in the proportion of individuals graduating in engineering.

In this section, we present an analysis that interprets the results in a human capital investment model including both opportunity costs and returns to education.[13] We present the estimated elasticity of educational choices to changes in returns following the oil discovery and compare the magnitudes to similar estimates proposed in the literature. Following the literature (see, e.g., Black et al., 2005a; Atkin, 2016), we assume that individuals choosing their education update their expectations of current and future earnings based on what they observe in the labor market. In Section 5.3, we explore alternative explanations for the changes in education, such as changes in public and private resources, finding no evidence of such mechanisms.

5.1 Empirical Model

We assign to every individual who is about to decide whether to enroll in education level *edu* in year y the average earnings of workers with that education level in year y in the respective labor market. Hence, all individuals who are born in the same cohort and the same labor market are assigned the same average earnings. We begin by estimating how these average earnings differ between oil regions, and how the difference evolves with time. First, we analyze the impact on the earnings of high school dropouts. An increase in the difference between oil regions would be evidence of an increase in the opportunity cost of completing high school. Next, to analyze changes in the returns to dropping out as compared to completing vocational high school, we investigate whether the earnings of vocational students also increased at the same time. Similarly, we estimate the evolution of returns to vocational high school as compared to completing university and the returns to engineering compared to other subjects. We estimate the parameters of the following regression:

$$log(earn_{c_i,llm_i,edu}) = \delta_0 + \sum_{c=1946}^{1984} \delta_c Cohort_{c_i} \times HighOil_i + \gamma_c Cohort_{c_i} + \gamma_{llm} LLM_i + \delta x_i + \varepsilon_{ic}. \quad (3)$$

[13]The estimated effects may also go through channels other than direct monetary incentives—the incentive may be through increased employment probability. For instance, risk-averse students may prefer education paths that lower the probability of having long spells of unemployment, even when wages for these types of paths are lower than the alternatives. If the oil sector buildup had an impact on employment probability, we may overestimate the impact of the monetary incentives. However, when we compare the employment rate of individuals of different ages and with different educations between oil regions, we find no changes related to the development phases (see Figure A7.). This analysis is explained in detail in Online Appendix C.

$log(earn_{c_i,llm_i,edu})$ is the log of the average earnings for working-age individuals with an educational level equal to edu, in the labor market where i is born and in the year the cohort c makes the decision about edu.[14] In particular, we focus on (i) the level of earnings for high school dropouts and (ii) vocational high school graduates, (iii) the ratio of earnings of university graduates to the earnings of vocational high school graduates, and (iv) the ratio of earnings for university graduates with an engineering degree to graduates in all other fields. All the variables on the right-hand side are the same as in Equation (1). Note that such a regression does not estimate the *causal* returns to education for individuals. Instead, δ_c captures the reduced form of log-difference in earnings between regions, i.e., the effect of the oil sector buildup on the average returns in the population.

In the next step, we link this effect to the educational choices of individuals. We estimate the following equation, where earnings are now an explanatory variable:

$$edu_{i,c,llm} = \beta_0 + \beta_1 log(earn_{c_i,llm_i,edu}) + \gamma_c + \gamma_{llm} + \beta_{\mathbf{x}}\mathbf{x_i} + e_{ic}. \quad (4)$$

To isolate the change in educational choices that is linked to the exogenous change in earnings due to the oil sector, we instrument the earnings variable, exploiting regional and cohort variation in the importance of the oil sector. Hence, Equation (3) becomes the first stage in the 2SLS procedure (for a similar approach see Black et al., 2005a; Morissette et al., 2015). β_1 is the percentage point increase in the proportion of individuals with education level edu when earnings (or returns) increase by 1%.

5.2 Estimates and Interpretation

In Figure 4, we report the estimates of the difference in earnings and returns over time between treated and control regions.[15] Panel (a) shows that earnings for high school dropouts increased significantly in high-oil areas in the early phase and remained higher throughout the sample, in line with the hypothesis of increased demand for low-skilled workers. Earnings for vocational high school graduates in Panel (b) followed a similar pattern. In contrast, returns to completing university in Panel (c) decreased significantly in oil-intensive regions for the years after the oil discovery and remained (insignificantly) lower throughout the sample. However, among university graduates in Panel (d), the returns for engineers increased slightly only in the middle of the mature phase.

[14] We look at earnings for individuals aged between 25 and 67 (the retirement age) in constant 1998 NOK (i.e. adjusted for inflation). We use an average over three years (moving window) to create a smoother indicator and because we are unable to exactly determine the year high school dropouts decide to leave school. However, the overall results are the same whether we use the 3-year average or the single-year measure.

[15] Since our earnings measure comes from the tax records starting in 1967 (see Section 3.4), we only observe the relevant earnings for a smaller number of cohorts. For example, for individuals selecting the type of high school, a decision taken at age 15, we only observe the relevant earnings measure starting from cohort 1952. Note, however, that the year 1967 is 3 years before the first oil discovery, so our data allow us to look at pre-trends. Note also that, given this data limitation, the dummy variable we exclude when estimating the parameters in Equation 3 for high school dropout and vocational high school is the one for cohorts born in 1952-1953.

Incorporating these results in a framework where both opportunity costs and returns to education shape individuals' choices, we can interpret the initial increase in the high school dropout rate as an immediate response to the increase in the opportunity cost of studying. On the other hand, the finding that high school dropouts return to school at a later point is likely driven by the gradual increase in earnings of vocational graduates. In a context where returning to school to complete vocational high school with a delay is feasible and at low costs, the decision to go back to school is justified for a rational forward-looking individual. Within the same framework, the increased earnings for vocational high school graduates in absolute terms imply that, by choosing an academic path, students would forego some earnings they would have otherwise earned had they not attended university (opportunity cost). Moreover, the wage increase for vocationally trained individuals compared to the earnings of university graduates implies that university graduates also face a relative decrease in returns. Thus, both forces push in the same direction, and explain the increase in vocational high school enrolment. The return to the education channel alone can easily rationalize the shift towards engineering, since the opportunity costs of any university degree are the same, while the returns differ in favor of engineering.

In Table 3, we report the estimates of the elasticity parameters. A 10% increase in earnings for high school dropouts at age 23 corresponds to an increase of approximately 2.9 percentage points in the proportion of students not completing high school (Column 1).[16] Although this parameter is not precisely estimated, this would imply an increase of 5.3% compared to the high school dropout rate in the same period in low-oil areas (Table A3), which is approximately the elasticity estimated by Black et al. (2005a) for the coal boom in Kentucky and Pennsylvania in the 1970s and 1980s. Hence, the magnitude of the initial reaction in terms of the high school dropout rate for a temporary coal price shock is similar to the effect of a more permanent shock in the setting of the buildup of the Norwegian oil sector. An explanation for the similarity of the effect might be students' inability to determine the temporary or permanent nature of an economic shock right from the beginning. This is supported by the evidence that students return to high school at a later point and indicates that the decision to drop out is likely taken without a careful long-term perspective. Overall, this suggests that the actual duration of the shock has little impact on the decision to drop out of high school when new opportunities open up.

The 10% decrease in returns to an academic degree as compared to vocational training leads to a decrease of 2.8 percentage points in university graduates, and, hence, an elasticity of about 6.3% (Column 2). Since our results suggest that the decision to attend university depends on the choices made at age 15 for the high school track, we also compute the same measure with returns at age 15 and estimate elasticities of a similar magnitude (Column 3). Overall, these numbers compare well to the estimates of Morissette et al. (2015) for an oil price increase in Canada. Nevertheless, it is larger than estimated for other types of temporary shocks: a back-of-the-envelope calculation

[16] Note that we only include the cohorts where we estimated an increase in the high school dropout rate in the sample estimating these parameters.

based on the results of Charles et al. (2018) indicates that for a 10% decrease in the returns to college following the housing boom in the US, college enrolment decreased by about 3.9%—about two-thirds of our estimate. Moreover, while Morissette et al. (2015) find that, following a resource price boom, students just postpone completion of university, Charles et al. (2018) find that the decrease in university enrolment from a housing boom is permanent. A comparison of our findings with the previous literature could thus indicate that the drop in university attainment is larger when the shock is more permanent or when the education decision is easily reversible.

The estimated elasticity of returns to engineering degrees is positive and large but is not precisely estimated. The imprecise estimates are likely a result of the small samples of engineering graduates, especially for cohorts born in the 1950s.[17]

5.3 Alternative Mechanisms

Other complementary explanations may also explain the difference in educational attainment across oil regions. In this section, we present additional evidence in support of the hypothesis that economic incentives are most important. In particular, we show that public investments were similar across oil regions, while private investments probably did not impact students' educational decisions.

Direct public investment in high-oil regions (potentially due to oil royalties) into educational institutions targeted to the oil sector could be an alternative explanation for our main findings. We argue that this mechanism is unlikely in our context. As mentioned, oil royalties are redistributed at the national level, and local authorities in high-oil regions have similar budgets as the rest of the country for schools, childcare, or higher education. In Figure A8, we plot the evolution of average school spending per capita at the municipality level in high- and low-oil regions from the 1960s to 2000. We find overlapping trends in spending for the two types of regions.[18] Regional differences in investment in higher education are also unlikely to drive our main findings. During the period we consider, new colleges were opened throughout the country with funding from the central government. However, the geographic location of these new colleges did not depend on the industry structure of the area; the expansion aimed to achieve an equal distribution of colleges throughout the country (see Carneiro et al., 2022). Hence, supply-side channels along this margin are not a central mechanism.

Oil-related changes in family resources or parental investment are another alternative potential mechanism. The oil discovery had an unambiguously positive effect on wages in the high-oil regions (Brunstad and Dyrstad, 1997), thus implying that children who grew up after the first oil discovery were likely to be living in more wealthy families. Higher family earnings are generally associated with higher education (Dahl and Lochner, 2012). Hence, it is possible that children from high-oil regions could modify their educational career as a consequence of higher household income, for example by completing high school instead of dropping out to work. We address this concern in all our regression

[17]Note that our estimates would predict an elasticity of approximately 40%. Given that the baseline numbers of engineering graduates are small, such a high elasticity is not implausible.

[18]The data are taken from the yearly municipal accounts.

analyses by controlling for parental earnings when the child was 17 years.[19] Moreover, Løken (2010) shows that the regional variation in the earnings increase due to the Norwegian oil boom did not affect the educational attainment of children growing up in more wealthy families. A reason for this finding may have been that richer parents in high-oil areas were not necessarily better educated and, hence, encouraged their children less to pursue an academic education path (Bütikofer et al., 2022).

6 Spillover Effects

So far, we have interpreted the change in economic incentives that followed the oil discoveries as a direct consequence of increased demand by the oil sector. However, it is possible that, as a consequence of the oil sector, demand for workers also increased in other sectors. For example, Black et al. (2005b) show that the 1970s coal boom generated an increase in employment not only in the mining sector but also in sectors supplying local goods such as retail and services. The increase in earnings of low-skilled and vocationally trained workers, as well as of engineers, may therefore also be a consequence of such spillover effects.

In this section, we exploit the richness of our data to investigate whether the type of education selected by students signals that they responded to changes in incentives from sectors other than the oil sector. We study this hypothesis in two steps. First, we exploit the fact that vocational education in Norway is very specific in terms of skill attainment, with students enrolling for training in one trade. We can thus establish whether the increase in vocational education is driven by more manufacturing-oriented or service-oriented tracks. Since the oil sector is mostly related to manufacturing-oriented education tracks (see, e.g., Brunstad and Dyrstad, 1997; Allcott and Keniston, 2017), evidence of an increase in the service tracks would point to spillover effects. Second, we turn our attention to women, who are less likely to be employed in the oil sector in Norway. We investigate their educational choices and see whether these choices can be interpreted as evidence of spillover effects.

Figure 5 plots the estimated coefficients of three different specifications of Equation 1, where we change the dependent variable: we use (i) a dummy for attending any vocational high school (the same as Panel (c) of Figure 2), (ii) a dummy for attending a manufacturing-oriented track at a vocational high school, and (iii) a dummy for attending a service-oriented track at a vocational high school. Panel (a) suggests that the entire difference in the proportion of men completing vocational high school is explained by the increase in manufacturing tracks.[20] This finding conforms that oil-related jobs are likely driving the growth in vocational high school education and suggest that the direct effect is more important for men than the spillover effects.

For women in Panel (b) of Figure 5, we observe no decrease in the percentage of individuals who are completing high school right after the first oil discovery as there is for men. However, similar to men, we find an increased probability of completing vocational high school in high-oil

[19] Note that controlling for parental earnings may not be sufficient if, for example, the functional form is not correctly specified or if the relevant variable is not current earnings but earnings expectations.

[20] Overall, in the period we consider, about 11% of the men enrolling in vocational high school do so in a service track.

areas for cohorts born after the mid-1960s. Compared to men, the effect size is smaller and the effect occurs with a delay. The increase among women comprises both manufacturing- and service-oriented tracks. This suggests that, if any spillover effects to other sectors are present, these are driven by women's educational choices. In addition, we do not find evidence of any increase in the high school dropout rate for women (Panel (a), Figure A9).

Spillovers on engineering education at the university level could result from oil revenues that are reinvested in public infrastructure in the high-oil areas, and therefore increase the demand for engineers outside of the oil sector in these regions (see, e.g., Mosquera, 2022). However, this type of explanation is not likely to hold in the Norwegian case, where the tax revenues from oil are equally redistributed across regions. The estimated increase in engineering degrees also applied to women (Panel (f), Figure A9). In addition, both men and women with engineering degrees at the university level are often employed in the oil sector. In the high-oil regions, this share is substantially higher than in the low-oil regions. In the high-oil regions, 28% of male engineers and 21% of female engineers born between 1960 and 1980 work in the oil sector by age 30, while the corresponding proportions are 9% and 7% in the low-oil regions. Overall, the small gender differences in oil sector employment among engineers indicate little spillover effects.

7 Conclusion

Given the importance of human capital for productivity, growth, and welfare, it is important to understand how large economic restructuring affects investment in education. In this paper, we show that permanent changes in the industrial composition have large impacts on the schooling choices of individuals mainly through changes in the labor market returns to different types of skills. We study the context of Norway and investigate the effect of oil discoveries in Norway in the late 1960s and the resulting buildup of the Norwegian oil sector on educational choices. This allows us to examine different stages of the economic change and to focus on different educational choices including high school completion, vocational vs. academic training, and field of study at the university level. We specify a dynamic difference-in-differences framework using cohorts who make educational decisions before the first oil discovery in Norway as the baseline and analyze the mechanisms driving the educational choices over the development of the industry and across educational margins.

We report six core findings. First, the rate of male students dropping out of high school increases significantly in the areas most affected by the oil sector in the immediate aftermath of the first oil discovery. We present empirical evidence that this is due to an increase in the opportunity cost of attending high school. However, this impact is only a short-run effect, and the regional differences in high school dropout rates even out for cohorts making educational decisions a decade after the first oil discovery. Second, we document that the increase in the high school dropout rate is a temporary effect even for the cohorts that initially quit high school as they complete high school as adults. Third, we document a large increase in the number of male students attending vocational secondary education and a decrease in students attending academic high school. This effect is largely driven by a large earnings increase for vocationally trained individuals relative to individuals with academic education. Hence, increased returns to vocational education likely offset the opportunity cost and thereby the incentive to not complete high school, and lead to a change in the type of education students choose. Fourth, we show that the proportion of university graduates in high-oil regions decreases permanently. Fifth, the decrease in university graduates is not uniform across fields of study. Engineering education is positively affected by oil discovery as returns to engineering degrees increase. Finally, we document that women, who are much less likely to work in the oil sector, are also more likely to enroll in vocational high school education as the industry matures which suggests that there are some spillover effects of the oil boom to the service sector.

In sum, a broad set of educational choices were affected by this economic restructuring and these choices were affected at different stages of the oil sector buildup. Hence, we provide some empirical evidence that the short-term and long-term consequences of large economic changes vary. Compared to previous studies that have predominantly focused on short-term changes (Morissette et al., 2015; Charles et al., 2018), the elasticity we estimate for the response to changes in earnings suggests that educational responses, especially in terms of high school specialization and university completion, are generally larger when a shock is more permanent. Most strikingly, we find that the elasticity is larger for educational decisions that are less easily reversible—a result that relates to previous studies showing that early educational tracking has negative consequences for longer-term outcomes (Pekkarinen et al., 2009).

Overall, this paper highlights the significant impact that large and long-term economic changes have on the human capital decisions of young individuals. The findings indicate that short-sighted educational decisions may lead to cohorts of workers who do not have general enough skills to respond to new economic challenges such as the current green transition. Potential areas for future research can include policies that allow for more reversible educational choices.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge, "When Should You Adjust Standard Errors for Clustering?," *The Quarterly Journal of Economics*, 10 2022, *138* (1), 1–35.

Allcott, Hunt and Daniel Keniston, "Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in Modern America," *Review of Economic Studies*, 2017, *85* (2), 695–731.

Atkin, David, "Endogenous Skill Acquisition and Export Manufacturing in Mexico," *American Economic Review*, 2016, *106* (8), 2046–85.

Bartik, Alexander, "Worker Adjustment to Changes in Labor Demand: Evidence from Longitudinal Census Data," 2017. Working paper.

Becker, Gary S., *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, Chicago: University of Chicago Press, 1964.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How Much Should We Trust Differences-in-Differences Estimates?," *The Quarterly Journal of Economics*, 2004, *119* (1), 249–275.

Bhuller, Manudeep, "Inndeling av Norge i Arbeidsmarkedsregioner," Notater 2009/24, Statistics Norway 2009.

Bjerkaker, Sturla, *Adult and continuing education in Norway*, W. Bertelsmann Verlag, 2016.

Black, Dan A., Terra G. McKinnish, and Seth G. Sanders, "Tight Labor Markets and the Demand for Education: Evidence from the Coal Boom and Bust," *Industrial and Labor Relations Review*, 2005, *59* (1), 3–16.

_ , Terra McKinnish, and Seth Sanders, "The Economic Impact Of The Coal Boom And Bust," *The Economic Journal*, 2005, *115* (503), 449–476.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes, "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education," *The Quarterly Journal of Economics*, 2005, *120* (2), 669–700.

Brunstad, Rolf Jens and Jan Morten Dyrstad, "Booming Sector and Wage Effects: An Empirical Analysis on Norwegian Data," *Oxford Economic Papers*, 1997, *49* (1), 89–103.

Bütikofer, Aline, Antonio Dalla-Zuanna, and Kjell G. Salvanes, "Breaking the Links: Natural Resource Booms and Intergenerational Mobility," *The Review of Economics and Statistics*, 2022, pp. 1–45.

Carneiro, Pedro, Kai Liu, and Kjell G Salvanes, "The Supply of Skill and Endogenous Technical Change: Evidence from a College Expansion Reform," *Journal of the European Economic Association*, 2022, *21* (1), 48–92.

Cascio, Elizabeth U and Ayushi Narayan, "Who Needs a Fracking Education? The Educational Response to Low-Skill-Biased Technological Change," *ILR Review*, 2022, *75* (1), 56–89.

Charles, Kerwin Kofi, Erik Hurst, and Matthew J Notowidigdo, "Housing booms and busts, labor market opportunities, and college attendance," *American Economic Review*, 2018, *108* (10), 2947–94.

Cust, James and Torfinn Harding, "Institutions and the Location of Oil Exploration," *Journal of the European Economic Association*, 2020, *18* (3), 1321–1350.

Dahl, Gordon B. and Lance Lochner, "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit," *American Economic Review*, 2012, *102* (5), 1927–56.

Ekeland, Anders, "Sysselsatte i Petroleumsnæringene og Relaterte Næringer 2014," Reports 2015/48, Statistics Norway 2015.

Emery, Herb, Ana Ferrer, and David Green, "Long Term Consequences of Natural Resource Booms for Human Capital Accumulation," *Industrial and Labor Relations Review*, 2012, *65* (3), 708–734.

Feyrer, James, Erin T. Mansur, and Bruce Sacerdote, "Geographic Dispersion of Economic Shocks: Evidence from the Fracking Revolution," *American Economic Review*, 2017, *107* (4), 1313–34.

Finansdepartementet, *Petroleumsvirksomhetens Plass i det Norske Samfunn*, Oslo: Norwegian Government, 1974.

Helle, Egil, *Norges Olje – de Første 20 Årene*, Oslo: Tiden, 1984.

Kovalenko, Alina, "Natural Resource Booms, Human Capital, and Earnings: Evidence from Linked Education and Employment Records," *American Economic Journal: Applied Economics*, April 2023, *15* (2), 184–217.

Løken, Katrine V., "Family Income and Children's Education: Using the Norwegian Oil Boom as a Natural Experiment," *Labour Economics*, 2010, *17* (1), 118 – 129.

Mincer, Jacob, "Investment in human capital and personal income distribution," *Journal of Political Economy*, 1958, *66* (4), 281–302.

— , "Human capital and economic growth," *Economics of Education Review*, 1984, *3* (3), 195–205.

Morissette, R., P. C. W. Chan, and Y. Lu, "Wages, Youth Employment, and School Enrollment: Recent Evidence from Increases in World Oil Prices," *Journal of Human Resources*, 2015, *50* (1), 222–253.

Mosquera, Roberto, "The long-term effect of resource booms on human capital," *Labour Economics*, 2022, *74*, 102090.

NOU, "Tilleggsberegninger — Ajourførte Pris-og Inntektsprognoser for 1972 og 1973 fra Det tekniske Beregningsutvalg," Rapport 1972/2, Regjeringen Korvald 1972.

OECD, *Education at a Glance* 2019.

Pekkarinen, Tuomas, Roope Uusitalo, and Sari Kerr, "School Tracking and Intergenerational Income Mobility: Evidence from the Finnish Comprehensive School Reform," *Journal of Public Economics*, 2009, *93* (7–8), 965 – 973.

Ryggvik, Helge, "A Short History of the Norwegian Oil Industry: From Protected National Champions to Internationally Competitive Multinationals," *Business History Review*, 2015, *89* (1), 3–41.

Weinstein, Russell, "Local labor markets and human capital investments," *Journal of Human Resources*, 2020, pp. 1119–10566R2.

Wooldridge, Jeffrey M, "Cluster-sample methods in applied econometrics," *American Economic Review*, 2003, *93* (2), 133–138.

8 Tables and Figures

Table 1: Estimated Difference in Education between Oil Regions by Development Phase, Men

	(1) No HS age 23	(2) No HS age 30	(3) Vocational high school	(4) Academic high school	(5) College	(6) Engineering (college)
Early	0.036***	0.034***	0.052***	-0.034***	0.016***	-0.035***
	(0.006)	(0.011)	(0.002)	(0.006)	(0.002)	(0.005)
Expansion	-0.056***	-0.064***	0.115***	-0.005	-0.016***	-0.064***
	(0.009)	(0.019)	(0.008)	(0.009)	(0.005)	(0.005)
Mature	-0.192***	-0.173***	0.174***	0.027***	0.022***	-0.116***
	(0.015)	(0.022)	(0.016)	(0.006)	(0.005)	(0.008)
High*Early	0.018**	0.002	0.003	0.000	-0.007	0.006
	(0.007)	(0.013)	(0.005)	(0.007)	(0.007)	(0.013)
High*Expansion	-0.014	-0.033*	0.056***	-0.021*	-0.011	0.027*
	(0.014)	(0.019)	(0.011)	(0.011)	(0.014)	(0.015)
High*Mature	-0.018	-0.024	0.050***	-0.028*	-0.026**	0.046***
	(0.018)	(0.022)	(0.016)	(0.015)	(0.012)	(0.014)
Observations	741,935	741,935	741,935	741,935	741,935	276,846
R-squared	0.200	0.175	0.054	0.162	0.179	0.016

Notes : Results for male students only. Each column comes from a separate regression, with the same explanatory variables and different outcome variables. The outcomes are dummy variables indicating high school completion in Columns (1) and (2), a vocational high school degree in Column (3), an academic high school degree in Column (4), a college degree in Column (5), and an engineering degree from a college in Column (6). In Column (6), we only include students who graduated from college. Where not specified, education levels are measured at age 30. Individuals are assigned to the sector development phase on the basis of the year at which they make the decision about an education level. We exclude the interaction between the high-oil region dummy and the dummy for the pre-oil cohort so that the pre-oil cohorts serve as the baseline. In the regression, we further control for cohort and labor market fixed effect, the number of siblings, the annual earnings of fathers in the year the individual turns 17, and the highest educational attainment of parents. We cluster the standard errors at the labor market level. *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Robustness Checks, Men

	(1) Baseline	(2) Municipality of residence	(3) Mother's birthplace	(4) Continuous measure	(5) Threshold down	(6) Threshold up
			Panel (a): High school dropout, age 23			
Early	0.018**	0.024***	0.025***	0.154	0.016**	0.020**
	(0.007)	(0.007)	(0.008)	(0.114)	(0.008)	(0.008)
Expansion	-0.014	-0.011	-0.013	-0.238	-0.016	-0.013
	(0.014)	(0.015)	(0.011)	(0.190)	(0.012)	(0.016)
Mature	-0.018	-0.011	-0.007	-0.257	-0.019	-0.014
	(0.018)	(0.018)	(0.015)	(0.229)	(0.018)	(0.020)
			Panel (b): High school dropout, age 30			
Early	0.002	0.002	0.008	-0.088	-0.004	0.007
	(0.013)	(0.013)	(0.014)	(0.227)	(0.014)	(0.013)
Expansion	-0.033*	-0.033	-0.024	-0.478	-0.035	-0.027
	(0.019)	(0.021)	(0.017)	(0.334)	(0.021)	(0.019)
Mature	-0.024	-0.020	-0.011	-0.352	-0.027	-0.020
	(0.022)	(0.022)	(0.018)	(0.306)	(0.023)	(0.021)
			Panel (c): Vocational high school			
Early	0.003	0.005	0.003	0.082	0.007	0.004
	(0.005)	(0.004)	(0.005)	(0.060)	(0.004)	(0.006)
Expansion	0.056***	0.057***	0.041***	0.671***	0.051***	0.060***
	(0.011)	(0.010)	(0.008)	(0.160)	(0.013)	(0.011)
Mature	0.050***	0.047***	0.027**	0.584**	0.044**	0.055***
	(0.016)	(0.016)	(0.012)	(0.223)	(0.019)	(0.015)
			Panel (d): Engineering			
Early	0.006	0.017	0.004	-0.044	-0.003	0.003
	(0.013)	(0.016)	(0.015)	(0.160)	(0.015)	(0.014)
Expansion	0.027*	0.044**	0.034**	0.162	0.018	0.023
	(0.015)	(0.017)	(0.014)	(0.196)	(0.017)	(0.018)
Mature	0.046***	0.064***	0.040***	0.366*	0.033	0.049***
	(0.014)	(0.015)	(0.012)	(0.203)	(0.020)	(0.011)

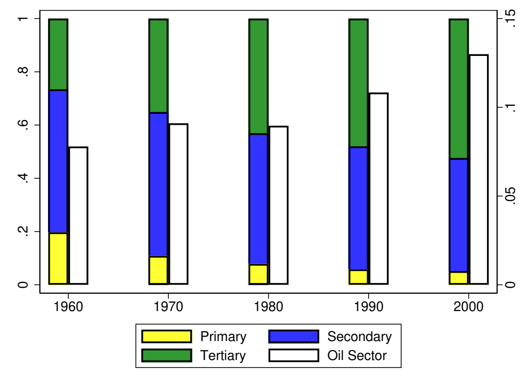
Notes : Each column and panel come from a separate estimate of Equation (2), where we change the definition of the measure of oil sector exposure. The outcomes are the same dummy variables as in Table 1. The estimates reported are for the parameters of the interaction between the dummy for the phase and the respective measure of oil sector exposure. In Column (1), we report our baseline estimates. We assign students to oil regions based on their place of residence at age 15 in Column (2) and on the birthplace of the mother in Column (3). In Column (4), we use a continuous measure of employment in the oil sector in 1980. We change the threshold for defining high- and low-oil regions to 6.5 and 9% in Column (5) and 8.5 and 11% in Column (6). We control for labor market fixed effect, the population in the municipality of birth in the year the person turns 15 (standardized to mean zero and standard deviation 1 in every cohort), the number of siblings, the birth order, the annual earnings of fathers in the year the individual turns 17, and the highest educational attainment of parents. We cluster the standard errors at the labor market level. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Elasticity of Education to Changes in Average Workers' Earnings, Men

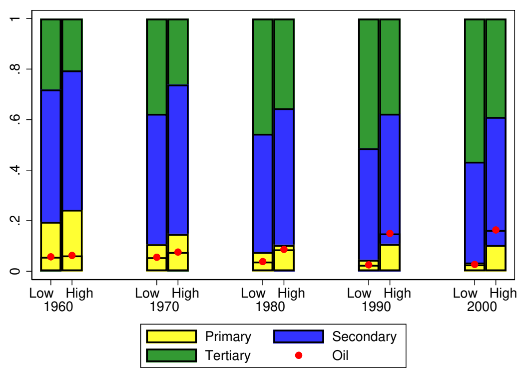
	(1) HS dropout (early cohorts)	(2) College (age 19)	(3) College (age 15)	(4) Engineering
OLS	0.105	0.141***	0.198***	-0.061
	(0.068)	(0.036)	(0.034)	(0.073)
2SLS	0.292	0.278**	0.268*	0.483
	(0.261)	(0.111)	(0.159)	(0.419)
Implied Elasticity	0.53	0.63	0.61	4.02
Observations	179,726	735,330	680,674	275,200

Notes : Each column comes from a separate regression, where educational outcomes are regressed on the earnings of workers with this level of education in the labor market. 2SLS specification instruments earnings with a dummy for being born in a high-oil region. Earnings variables are three-year averages of the gross earnings of individuals between ages 25 and 67 in the labor market of birth in the year individuals choose the specific education. All regressions control for cohort and labor market fixed effects, population, number of siblings, birth order, education of parents, and earnings of fathers. Standard errors clustered at the labor market level. Implied elasticity is the percentage change in the proportion of men with each education implied by the 2SLS parameter compared to the proportion of men with the same education in the low-oil areas during the mature phase (early phase for high school dropout). *** $p<0.01$, ** $p<0.05$, * $p<0.1$

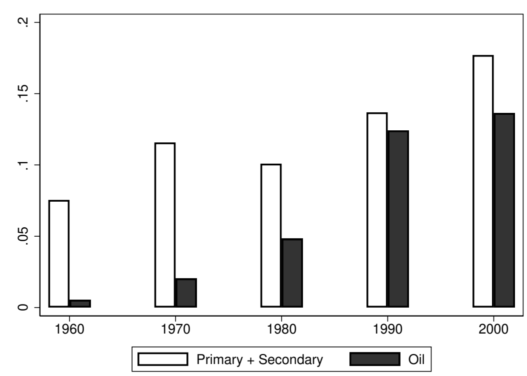
Figure 1: Sectoral Employment and Education for Men, Age 30-35



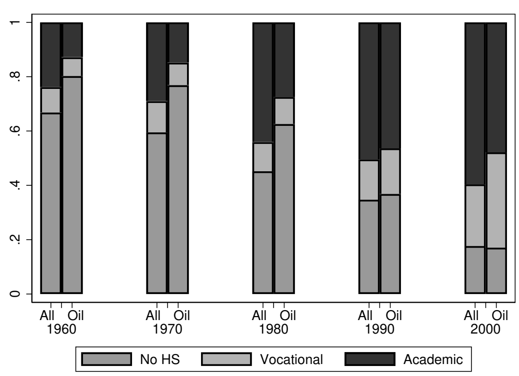
(a) Proportions Employed in Each Sector



(b) Sectors of Employment by Oil Area



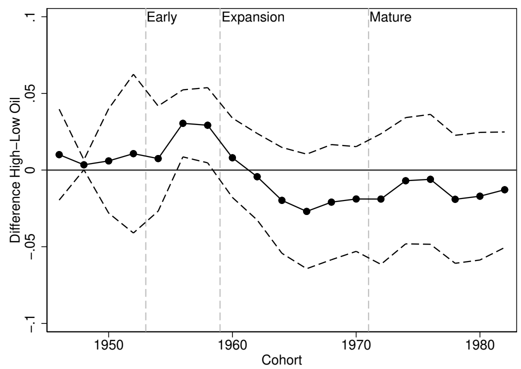
(c) Differences in Sectors between Oil Area



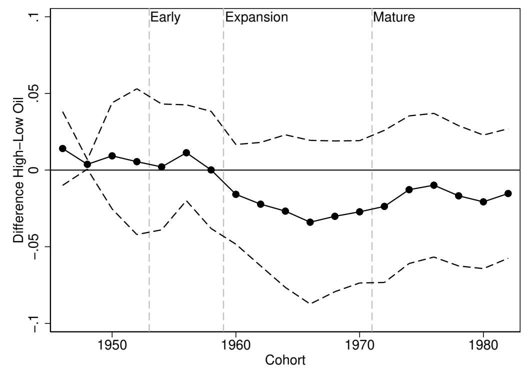
(d) Education Levels by Sector of Employment

Notes : Panel (a) shows the proportion of male workers between age 30 and 35 employed in the primary, secondary and tertiary sector (coloured bar, left-axis), and the proportion employed in the oil sector (right axis). Panel (b) shows the same proportions, separately for low-oil ("Low" in the label) and high-oil ("High") regions. Panel (c) shows the difference between high- and low-oil regions in the proportion of workers employed in the primary and secondary sectors ("Primary+Secondary" in the label) and the difference in the proportion of workers employed in the oil sector ("Oil" in the label). Panel (d) reports the proportion of workers who dropped out of high school ("No HS"), who chose a vocational track, and who chose an academic track, separately for the overall population ("All") and for workers employed in the oil sector ("Oil").

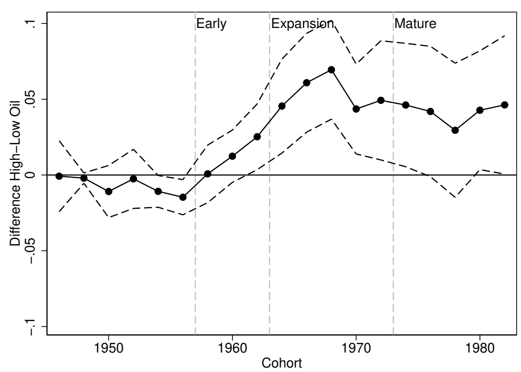
Figure 2: Estimated Difference Between High- and Low-Oil Regions by Educational Level, Men



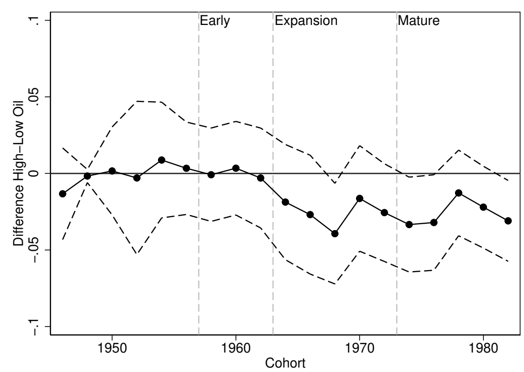
(a) High school dropout, age 23



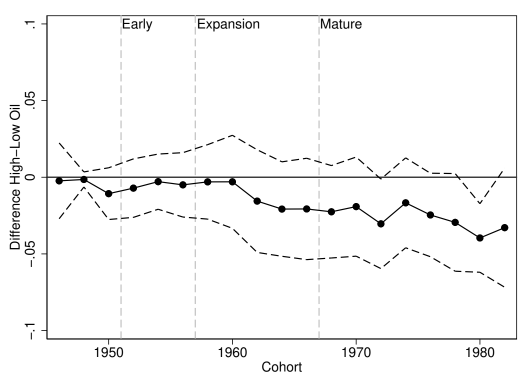
(b) High school dropout, age 30



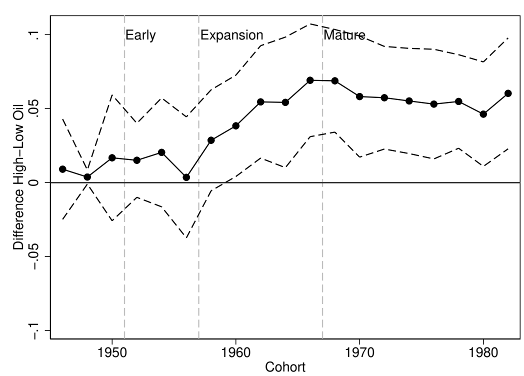
(c) Vocational high school



(d) Academic high school



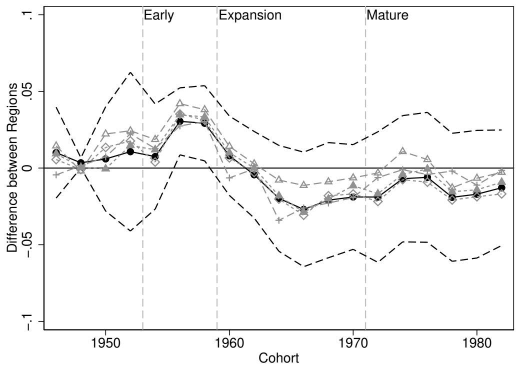
(e) College



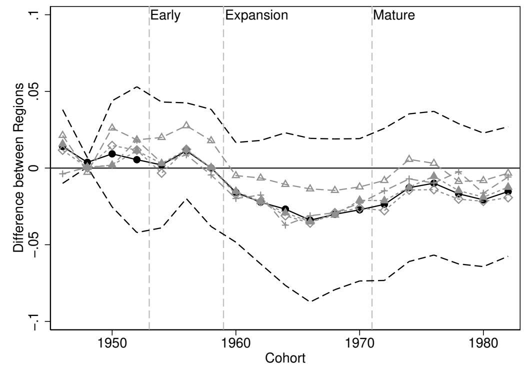
(f) Engineering, college graduates

Notes : The sample includes male students only. The graphs plot the estimates of the parameters from Equation (1) for the interaction terms between a dummy for being born in the high-oil region and a dummy for being born in each birth cohort. The dependent variables are different educational outcomes. Panel (f) includes only students who graduated from college. Where not specified, education levels are measured at the age of 30 years. We exclude the interaction between the high-oil region dummy and the dummy for the pre-oil cohort so that the pre-oil cohorts serve as the baseline. We control for cohort and labor market fixed effect, the population in the municipality of birth in the year the person turns 15 (standardized to mean zero and standard deviation 1 in every cohort), the number of siblings, the birth order, the annual earnings of fathers in the year the individual turns 17, and the highest educational attainment of parents. We cluster the standard errors at the labor market level. Dashed lines show 95% confidence intervals for the estimated parameters. Vertical lines separate the development phases during which the cohorts make different education decisions.

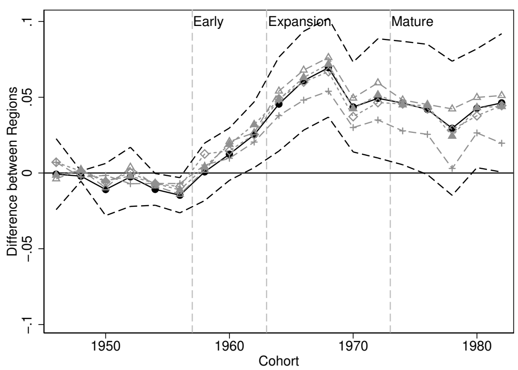
Figure 3: Robustness Checks



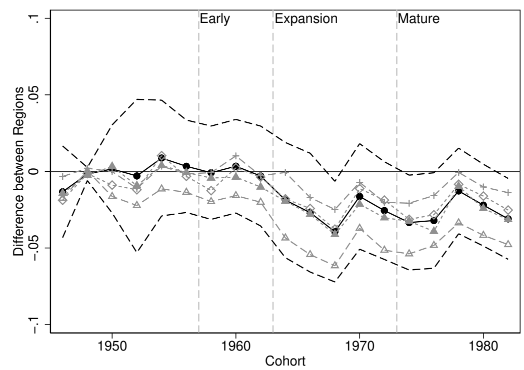
(a) High School Dropout, Age 23



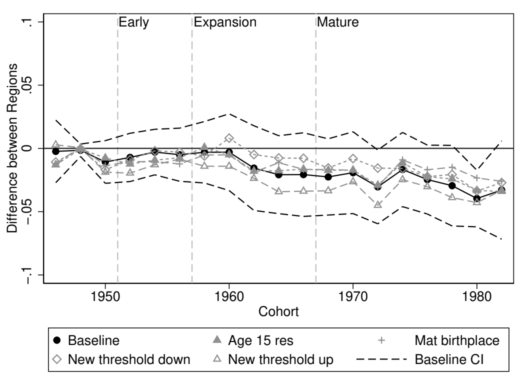
(b) High School Dropout, Age 30



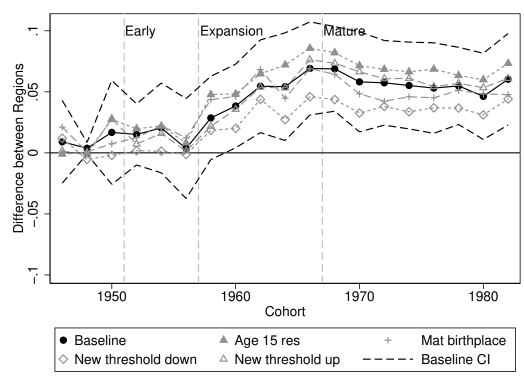
(c) Vocational High School



(d) Academic High School



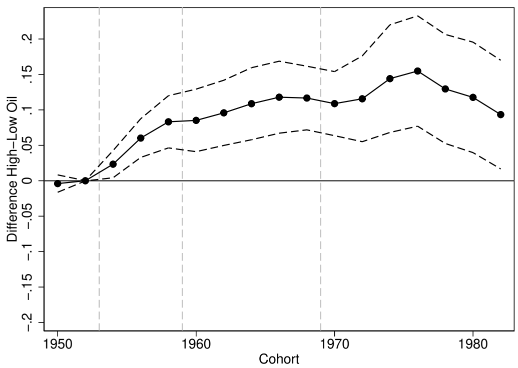
(e) University



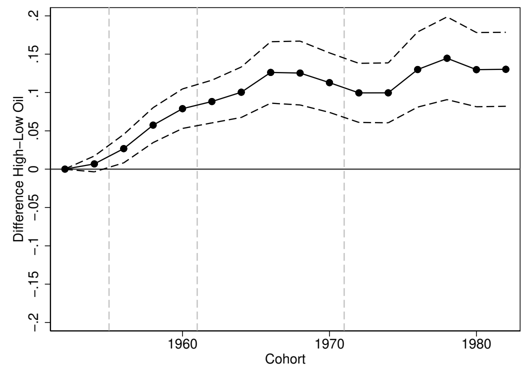
(f) Engineering, University Graduates

Notes : Estimates of the difference between high- and low-oil regions from Equation (1), using different definitions of oil regions. "Baseline" assigns students on the basis of the place of birth and defines low-oil areas as all LLMs where oil employment among workers in 1980 was $< 7.5\%$ and high-oil areas where oil employment was $> 10\%$ (as in Figure 2). "Age 15 res" assigns students to the place where they live in the year they turn 15 years. "Mat. birthplace" assigns students to the mother's place of birth. "New threshold down" defines low-oil areas as labor markets where $< 6.5\%$ and high-oil as those where $> 9\%$ of workers were in the oil sector in 1980."New threshold up" uses $< 8.5\%$ and $> 11\%$ as thresholds. The black dashed line is the 95% confidence interval for the "Baseline" estimates, where standard errors are clustered at the labor market level.

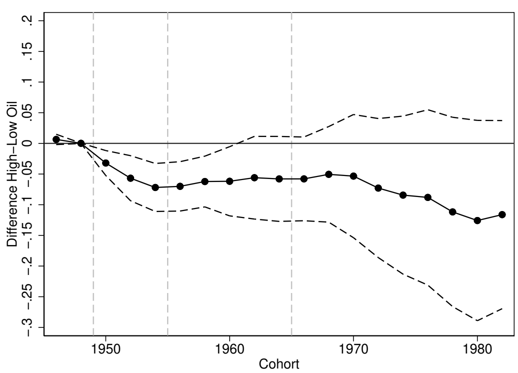
Figure 4: Estimated Returns to Different Types of Education by Regions



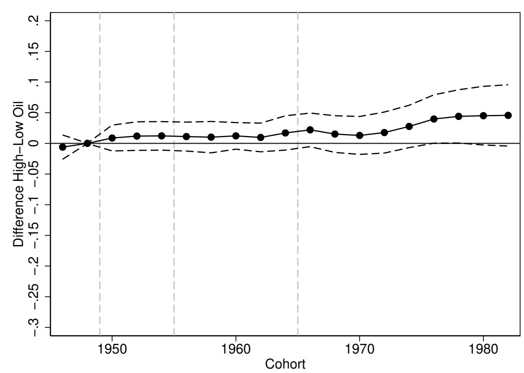
(a) HS dropouts



(b) Vocational high school



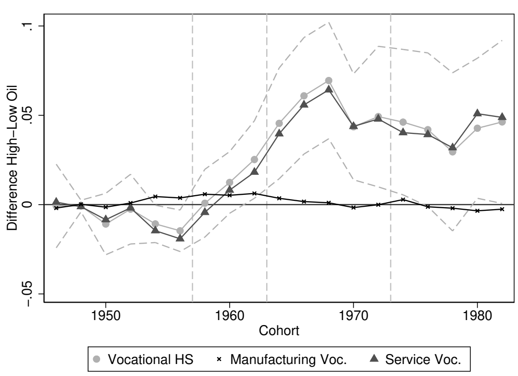
(c) Returns to college vs. vocational



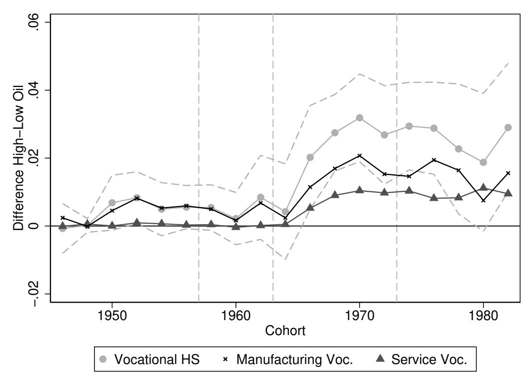
(d) Engineering vs. other college

Notes : Each figure plots the estimated difference between oil regions in the average earnings of workers with specific education levels in the year each cohort on the x-axis makes their educational decision, exploiting the framework in Equation (3). All dependent variables are expressed in log terms. Panel (a) focuses on high school dropouts, Panel (b) examines individuals with vocational high school degrees, Panel (c) compares individuals who graduated from a university and those who completed vocational high school, and Panel (d) compares individuals with an engineering degree and all other college graduates.

Figure 5: Regional Difference in the Types of Vocational Education



(a) Vocational HS, men



(b) Vocational HS, women

Notes : The graphs plot the estimates of the parameters from Equation (1) for the interaction terms between a dummy for being born in the high-oil region and a dummy for being born in each birth cohort, where the dependent variables are indicators for different vocational educations. "Vocational HS" are the estimates for all vocational high school tracks, jointly. "Manufacture" includes manufacture-oriented vocational high school tracks only, and includes mechanics (electrical, electronic, mechanic, and machine subjects), building, and extraction. "Service" includes service-oriented courses selected within vocational high school and includes business and accounting, tourism, personal care, wholesale, retail sale, and social service. We exclude the interaction between the high-oil region dummy and the dummy for the pre-oil cohort so that the pre-oil cohorts serve as the baseline. The regression further controls for cohort and local labor market fixed effect, the population in the municipality of birth in the year the person turns 15 (standardized to mean zero and standard deviation 1 in every cohort), the number of siblings, and the birth order, the annual earnings of fathers in the year the individual turns 17 and the highest educational attainment of parents. Standard errors are clustered at the labor market level. Dashed lines show 95% confidence intervals for the estimated parameters. Vertical lines separate the development phase during which the cohorts make different education decisions.

ONLINE APPENDIX

A The Evolution of the Norwegian Oil Industry

Oil exploration in the Norwegian Sea started in 1965 when the Norwegian state issued production licenses for offshore areas around the southwestern tip of Norway. As Norway lacked the necessary knowledge in the 1960s, all the relevant equipment to drill oil wells was imported from overseas, mostly from the US, and this activity had little impact on the mainland economic environment. In addition, the first attempts to find oil or gas reserves failed. It was only after four years of exploration that Ekofisk, one of the largest offshore oilfields ever found, was discovered on December 23, 1969, below the sea around 300 km southwest of the city of Stavanger. Production from Ekofisk commenced in 1971 (see, e.g., Helle, 1984).

In 1972, the Norwegian parliament voted to increase regulation of oil exploration and to develop new knowledge and industries based on petroleum (Finansdepartementet, 1974). To safeguard national interests in this new sector, Statoil, a state-owned oil company, was formed in 1972 and the government decided that petroleum from the Norwegian continental shelf must only be landed in Norway (NOU, 1972). The development of Statoil into a fully operative company took several years, and the initial activities directly linked to the exploration and exploitation of oilfields were therefore dominated by foreign companies.[21] In particular, the core oil drilling technologies were imported and installed by foreign companies. However, already from the early 1970s, some Norwegian companies were directly involved in the supply industry. These companies include shipping, shipbuilding, and marine engineering firms with comparative advantage in the development and construction of vessels operating in harsh conditions in the North Sea. Hence, during this early phase, which lasted for the first half of the 1970s, very few Norwegian workers benefited from the new sector, and activities were concentrated in the coastal area close to Ekofisk and the areas where the shipping industry was more prevalent.[22]

Starting from the second half of the 1970s, under pressure from calls for a more protectionist approach, the Norwegian government made it a goal to develop Norwegian skills in all the sectors involved in the oil industry.[23] As an example of the policies that followed this decision, companies had to declare the proportion of Norwegian involvement and only oil companies that contributed to the "Norwegianization" of the industry would be awarded new licenses. In addition, Statoil and the Norwegian government played an active part in developing local competencies, by actively participating in building the platforms and by contributing to the creation of a new engineering firm in 1977. The dominant position of Statoil, with its headquarters in Stavanger, contributed to the concentration of oil activity in the area.

[21] For a detailed description of the history of Statoil and the Norwegian oil industry, see Ryggvik (2015).

[22] To quantify, Ryggvik (2015) reports that in 1975 the local content of all offshore activities was still only around 28%.

[23] A more protectionist approach was advocated following the crisis in the shipbuilding industry and there was an overall international trend towards more protectionist policies, for example in the UK.

These protectionist policies were continued by different governments (both social democratic and more conservative) and resulted in an expansion phase of the oil industry, which ended in 1986 when taxes for foreign companies were significantly reduced and the share of foreign participation in the new licensing rounds increased significantly. This was partly a reaction to low oil prices and fears that foreign companies would leave the country, but it was also based on the presumption that the Norwegian oil industry had reached a mature phase, where Statoil was able to compete internationally (Ryggvik, 2015). Despite the more competitive environment, Norwegian firms supplying the oil industry retained their large share of the market in the mature phase. In addition, the competencies these firms had developed in deep sea drilling also proved to be valuable in other markets, and several companies expanded internationally. Nevertheless, the initial geographical distribution of the companies involved in the oil industry still shaped the industry's later geographic distribution, as companies founded during the expansion phase are still among the biggest oil companies in the country.

B Education Data and the Norwegian Education System

We used the educational database provided by Statistics Norway to measure educational attainment. Since 1974, educational attainment is reported annually by the educational institutions directly to Statistics Norway, thereby minimizing any measurement error. For individuals who completed their education before 1974, we use self-reported information from the 1960 and 1970 censuses, which are considered to be very accurate (see, e.g., Black et al., 2005c). We define three education levels: dropped out or not completed high school, completed high school, and completed higher education (college and university). We also observe whether an individual was enrolled in a vocational or an academic high school track, and we can further analyze which specific field within the vocational track students choose. During the period we investigate, there were nine distinct programs teaching students vocational qualifications including "health and child development", "restaurant and food", or "construction". We also categorize higher education into different fields of study. In particular, we looked separately at engineering degrees, thereby isolating those more involved in the oil industry.

Earnings data come from administrative tax records beginning in 1967 and are pre-tax earnings. In addition to labor market income from wages and self-employment, the earnings measure includes transfers from the social security system, such as unemployment benefits, sickness benefits, and parental leave benefits. The earnings measure is not top-coded and includes labor earnings expressed in constant 1998 Norwegian kroner (hence adjusted for inflation).[24]

[24] Since our earnings measure comes from the tax records starting in 1967, when we estimate the earnings elasticity of educational choices, we only observe the relevant earnings for a smaller number of cohorts since the earnings records start in 1967. For example, for individuals selecting the type of high school, a decision taken at age 15, we observe the earnings of vocationally trained workers only starting from the 1952 cohort and not from the 1947 birth cohort which would be ideal for our setting. However, 1967 is three years before the oil boom started, so our data enable us to assess pre-trends.

C Impact of the Employment Rate on Educational Decisions

Following the literature estimating elasticities, we interpret the changes in educational decisions as a response to monetary incentives. However, the oil boom may have had other types of impacts, such as an increased number of available jobs for different types of education. To assess the importance of this labor demand channel, we investigate whether differences in the employment rate changed between regions and whether the employment rate at the time when the school decision was made had an impact on the decision.

Figure A10 reports the evolution of the national unemployment rate for male workers together with the measure we use for non-employment (the reciprocal of the employment rate), corresponding to individuals reporting zero earnings. We rely on this non-employment measure because it allows us to separate it by education level and local labor markets. The two measures display similar patterns and suggest that the early phase of the oil sector buildup coincided with a period of high employment throughout the country; the unemployment rate was on average around 1% between 1970 and 1980. The unemployment rate increased in the early 1980s and during the 1990s as Norway went through periods of recession, but it never exceeded 6%. However, it is possible that, especially during recessions, some sectors coped better than others, creating differences in the employment rate depending on the skills required in the sector in question.

We then look at the differences in the employment rate between regions, separately for individuals with different types of education. We also study the employment rate for different age groups, which can offer alternative explanations for behavior relating to schooling choices. For example, if increased employment can explain the increased opportunity cost of schooling and the increased high school dropout rate, we should observe an increase in the employment rate for the age group 18-20 years. If, instead, the returns to schooling are increased by the fact that finding a job is easier in the long run, this should be reflected in an increased employment rate at a later age (see, e.g., Cascio and Narayan, 2022).

Figure A7 plots the raw difference in the employment rate between high-oil and low-oil regions for different age groups. Panel (a) shows the evolution of this difference over time for individuals who dropped out of high school, while Panel (b) displays the graph for individuals with a vocational high school qualification. Each dot in the graphs represents the difference in the year the respective cohort (reported on the x-axis) decides whether to drop out of high school (Panel (a)) or whether to enroll in vocational high school (Panel (b)). The dashed vertical lines separate the cohorts deciding about the education level during each development phase. Overall, the graphs show a very constant difference between regions over time, which is mostly close to zero. This evidence supports the view that the creation of job opportunities for non-employed individuals is not a relevant force driving oil industry-related differences in educational choices between regions.

Tables and Figures

Table A1: Regional Difference in Degree Completion, Aggregated on Labor Market Level

	(1) No HS, age 23	(2) No HS, age 30	(3) Vocational high school	(4) Academic high school	(5) College	(6) Engineering (college)
High * 1946	0.005	0.008	-0.015	0.007	0.004	-0.006
	(0.014)	(0.016)	(0.012)	(0.015)	(0.011)	(0.017)
High * 1950	0.008	0.015	-0.019*	0.004	-0.014	0.016
	(0.020)	(0.017)	(0.011)	(0.016)	(0.012)	(0.016)
High * 1952	0.018	0.015	-0.015	-0.000	-0.005	0.013
	(0.028)	(0.023)	(0.011)	(0.022)	(0.012)	(0.015)
High * 1954	0.017	0.017	-0.022*	0.006	-0.005	0.013
	(0.015)	(0.014)	(0.012)	(0.015)	(0.012)	(0.015)
High * 1956	0.035***	0.022*	-0.025**	0.003	-0.005	0.002
	(0.013)	(0.013)	(0.012)	(0.013)	(0.013)	(0.014)
High * 1958	0.034***	0.009	-0.004	-0.005	-0.008	0.025*
	(0.013)	(0.012)	(0.013)	(0.014)	(0.012)	(0.014)
High * 1960	0.016	-0.003	0.006	-0.004	-0.010	0.032**
	(0.014)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
High * 1962	0.003	-0.007	0.017	-0.011	-0.021	0.049***
	(0.014)	(0.013)	(0.012)	(0.015)	(0.014)	(0.016)
High * 1964	-0.010	-0.009	0.032***	-0.023	-0.023*	0.044**
	(0.013)	(0.014)	(0.012)	(0.014)	(0.013)	(0.019)
High * 1966	-0.019	-0.016	0.050***	-0.034**	-0.025*	0.059***
	(0.013)	(0.015)	(0.012)	(0.015)	(0.014)	(0.012)
High * 1968	-0.012	-0.010	0.054***	-0.044***	-0.026**	0.059***
	(0.013)	(0.013)	(0.012)	(0.013)	(0.012)	(0.012)
High * 1970	-0.007	-0.004	0.025**	-0.021	-0.023*	0.047***
	(0.013)	(0.013)	(0.012)	(0.014)	(0.013)	(0.013)
High * 1972	-0.009	0.001	0.029**	-0.030**	-0.030**	0.048***
	(0.014)	(0.013)	(0.012)	(0.014)	(0.012)	(0.014)
High * 1974	0.003	0.012	0.025**	-0.038**	-0.018	0.043***
	(0.014)	(0.013)	(0.013)	(0.016)	(0.012)	(0.012)
High * 1976	0.009	0.017	0.023*	-0.040**	-0.029**	0.044***
	(0.014)	(0.014)	(0.013)	(0.018)	(0.013)	(0.013)
High * 1978	-0.002	0.014	0.009	-0.023	-0.032***	0.043***
	(0.014)	(0.015)	(0.013)	(0.014)	(0.012)	(0.015)
High * 1980	-0.002	0.009	0.022*	-0.031*	-0.041***	0.035***
	(0.014)	(0.015)	(0.012)	(0.016)	(0.014)	(0.013)
High * 1982	0.003	0.013	0.026**	-0.039***	-0.034**	0.054***
	(0.013)	(0.014)	(0.013)	(0.014)	(0.013)	(0.012)
Constant	0.463***	0.693***	0.203**	0.334***	0.196***	0.228***
	(0.087)	(0.067)	(0.080)	(0.098)	(0.069)	(0.046)
N	800	800	800	800	800	800

Notes : Results for male students only. Where not specified, education levels are considered at age 30. Each column comes from a separate regression, where we vary the dependent variable. Each row is the estimate of the parameter on the interaction term between a dummy for being born in the high-oil region and a dummy for being born in each couple of cohorts from Equation (1). We exclude the interaction between the high-oil region dummy and the dummy for being born in the 1948 and 1949 cohorts. Observations are collapsed at the labor market by cohort level. The estimates are weighted by population size. Newey-West standard errors allow for autocorrelation of order 1 within the labor market in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

Table A2: Cohorts by Educational Choices, Development Phase, and Oil Region

	Completing high school *(Choice at age 16-18)*			Type of high school *(Choice at age 15)*			College *(Choice at age 19)*		
Men									
	Cohorts	Size Low Oil	Size High Oil	Cohorts	Size Low Oil	Size High Oil	Cohorts	Size Low Oil	Size High Oil
Pre-Oil	1946-1952	108,028	24,116	1946-1956	154,444	35,007	1946-1950	78,264	16,992
Early	1953-1958	94,520	21,996	1956-1960	79,954	18,598	1951-1955	76,180	18,015
Expansion	1959-1968	180,198	44,935	1961-1970	188,906	48,395	1956-1965	167,524	39,956
Mature	1969-1984	272,758	81,033	1971-1984	232,200	70,080	1966-1984	333,536	97,117
Women									
	Cohorts	Size Low Oil	Size High Oil	Cohorts	Size Low Oil	Size High Oil	Cohorts	Size Low Oil	Size High Oil
Pre-Oil	1946-1952	102,939	22,917	1946-1956	147,830	33,542	1946-1950	74,484	16,333
Early	1953-1958	91,528	21,367	1956-1960	77,228	17,862	1951-1955	73,346	17,209
Expansion	1959-1968	171,924	43,213	1961-1970	179,817	46,590	1956-1965	160,986	38,512
Mature	1969-1984	260,914	77,271	1971-1984	222,430	66,774	1966-1984	318,489	92,714

Notes : Different cohorts are assigned to different phases for each educational choice, because students decide about which education to obtain at different ages.

Table A3: Share of Individuals Completing Education by Development Phase and Oil Region

	No HS, age 23		No HS, age 30		Vocational high school	
Men	Low	High	Low	High	Low	High
Pre-Oil	0.56	0.55	0.46	0.46	0.09	0.08
Early	0.53	0.55	0.45	0.46	0.13	0.13
Expansion	0.40	0.39	0.32	0.29	0.18	0.24
Mature	0.17	0.15	0.13	0.11	0.21	0.27
Women	Low	High	Low	High	Low	High
Pre-Oil	0.56	0.54	0.49	0.48	0.02	0.02
Early	0.53	0.54	0.50	0.51	0.03	0.03
Expansion	0.34	0.34	0.29	0.30	0.09	0.10
Mature	0.13	0.13	0.10	0.10	0.11	0.13

	Academic high school		College		Engineering (college)	
Men	Low	High	Low	High	Low	High
Pre-Oil	0.45	0.45	0.23	0.24	0.25	0.24
Early	0.46	0.45	0.29	0.28	0.21	0.21
Expansion	0.54	0.50	0.29	0.28	0.18	0.20
Mature	0.67	0.63	0.44	0.40	0.12	0.16
Women	Low	High	Low	High	Low	High
Pre-Oil	0.48	0.48	0.17	0.18	0.01	0.01
Early	0.53	0.52	0.25	0.23	0.02	0.01
Expansion	0.66	0.64	0.30	0.29	0.03	0.03
Mature	0.81	0.79	0.56	0.54	0.02	0.02

Notes : Where not specified, education level is observed at age 30. "No HS" is the proportion of students who did not complete high school. The last column of the last panel "Engineering (college)" is the proportion of students who graduated with an engineering degree only among students graduating from college.

Table A4: Regional Difference in Degree Completion

	(1) No HS, age 23	(2) No HS, age 30	(3) Vocational high school	(4) Academic high school	(5) College	(6) Engineering (college)
High * 1946	0.010	0.014	-0.001	-0.013	-0.002	0.009
	(0.015)	(0.012)	(0.012)	(0.015)	(0.012)	(0.017)
High * 1950	0.006	0.009	-0.011	0.002	-0.011	0.017
	(0.017)	(0.017)	(0.009)	(0.014)	(0.008)	(0.021)
High * 1952	0.011	0.005	-0.003	-0.003	-0.007	0.015
	(0.026)	(0.023)	(0.010)	(0.025)	(0.009)	(0.012)
High * 1954	0.008	0.002	-0.011**	0.009	-0.003	0.021
	(0.017)	(0.020)	(0.005)	(0.019)	(0.009)	(0.018)
High * 1956	0.030***	0.011	-0.015**	0.003	-0.005	0.004
	(0.011)	(0.015)	(0.006)	(0.015)	(0.010)	(0.020)
High * 1958	0.029**	0.000	0.001	-0.001	-0.003	0.029*
	(0.012)	(0.019)	(0.009)	(0.015)	(0.012)	(0.017)
High * 1960	0.008	-0.016	0.012	0.003	-0.003	0.038**
	(0.013)	(0.016)	(0.009)	(0.015)	(0.015)	(0.017)
High * 1962	-0.004	-0.022	0.025**	-0.003	-0.016	0.055***
	(0.014)	(0.020)	(0.011)	(0.016)	(0.017)	(0.019)
High * 1964	-0.020	-0.027	0.046***	-0.019	-0.021	0.054**
	(0.017)	(0.025)	(0.015)	(0.019)	(0.015)	(0.022)
High * 1966	-0.027	-0.034	0.061***	-0.027	-0.021	0.069***
	(0.018)	(0.026)	(0.016)	(0.019)	(0.016)	(0.019)
High * 1968	-0.021	-0.030	0.069***	-0.039**	-0.023	0.069***
	(0.019)	(0.024)	(0.016)	(0.016)	(0.015)	(0.017)
High * 1970	-0.019	-0.027	0.044***	-0.016	-0.019	0.058***
	(0.017)	(0.023)	(0.015)	(0.017)	(0.016)	(0.020)
High * 1972	-0.019	-0.024	0.049**	-0.026	-0.030**	0.057***
	(0.021)	(0.025)	(0.019)	(0.016)	(0.014)	(0.017)
High * 1974	-0.007	-0.013	0.046**	-0.033**	-0.017	0.055***
	(0.020)	(0.024)	(0.020)	(0.015)	(0.014)	(0.018)
High * 1976	-0.006	-0.010	0.042*	-0.032**	-0.025*	0.053***
	(0.021)	(0.023)	(0.021)	(0.015)	(0.013)	(0.018)
High * 1978	-0.019	-0.017	0.030	-0.013	-0.029*	0.055***
	(0.021)	(0.023)	(0.022)	(0.014)	(0.016)	(0.016)
High * 1980	-0.017	-0.021	0.043**	-0.022	-0.040***	0.046**
	(0.021)	(0.022)	(0.019)	(0.013)	(0.011)	(0.017)
High * 1982	-0.013	-0.015	0.046**	-0.031**	-0.033*	0.060***
	(0.019)	(0.021)	(0.023)	(0.013)	(0.019)	(0.019)
Constant	0.337***	0.371***	0.294***	0.369***	0.169***	0.137***
	(0.015)	(0.018)	(0.012)	(0.013)	(0.013)	(0.005)
N	741,935	741,935	741,935	741,935	741,935	276,846

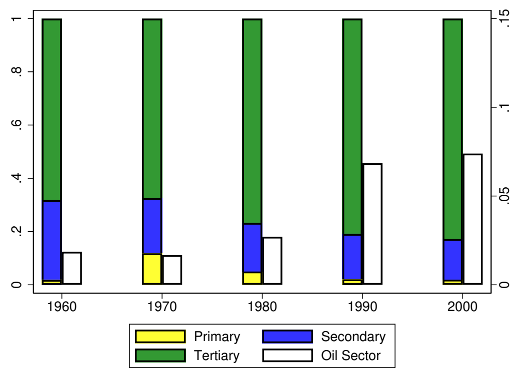
Notes : Results for male students only. Where not specified, education levels are considered at age 30. Each column comes from a separate regression, where we vary the dependent variable. Each row is the estimate of the parameter on the interaction term between a dummy for being born in the high-oil region and a dummy for being born in each couple of cohorts from Equation (1). We exclude the interaction between the high-oil region dummy and the dummy for being born in the 1948 and 1949 cohorts. We control for cohort and labor market fixed effects, the population in the municipality of birth in the year the person turns 15 (standardized to mean of zero and standard deviation 1 in every cohort), the number of siblings, the birth order, the annual earnings of fathers in the year the individual turns 17, and the highest educational attainment of parents. We cluster the standard errors at the labor market level. *** p<0.01, ** p<0.05, * p<0.1.

Table A5: Estimated Difference in Education between Oil Regions by Phase, Aggregated on Labor Market Level

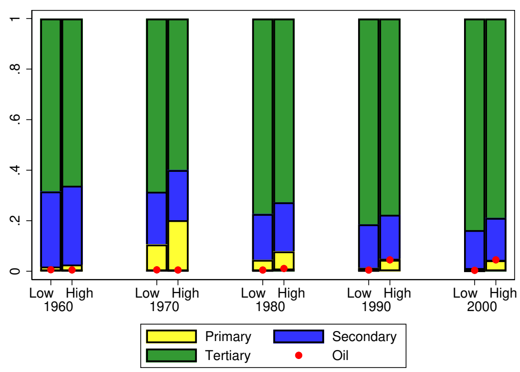
	(1) No HS, age 23	(2) No HS, age 30	(3) Vocational high school	(4) Academic high school	(5) College	(6) Engineering (College)
Early	0.043***	0.042**	0.094***	-0.103***	0.011	-0.045***
	(0.016)	(0.017)	(0.019)	(0.022)	(0.013)	(0.013)
Expansion	-0.034	-0.045	0.176***	-0.117***	-0.022	-0.088***
	(0.027)	(0.030)	(0.029)	(0.032)	(0.028)	(0.023)
Mature	-0.159***	-0.173***	0.275***	-0.116**	0.034	-0.139***
	(0.035)	(0.041)	(0.043)	(0.048)	(0.039)	(0.029)
High*Early	0.019	0.005	0.003	-0.001	-0.005	0.002
	(0.018)	(0.017)	(0.012)	(0.011)	(0.015)	(0.015)
High*Expansion	-0.015	-0.025	0.052***	-0.025**	-0.007	0.017
	(0.015)	(0.016)	(0.011)	(0.012)	(0.016)	(0.017)
High*Mature	-0.017	-0.010	0.039***	-0.027*	-0.017	0.032**
	(0.017)	(0.017)	(0.012)	(0.014)	(0.015)	(0.013)
Constant	0.481**	0.575**	0.638***	-0.373	0.107	0.177
	(0.223)	(0.256)	(0.202)	(0.235)	(0.228)	(0.134)
Observations	160	160	160	160	160	160
R-squared	0.994	0.990	0.951	0.973	0.982	0.939

Notes : Results for male students only. Each column comes from a separate regression, with the same explanatory variables and different outcome variables. In Column (6), we only include students who graduated from college. Where not specified, education levels are measured at age 30. Individuals are assigned to the development phase on the basis of the year at which they make the decision about an education level. We exclude the interaction between the high-oil region dummy and the dummy pre-oil cohort dummy so that the difference between oil regions for pre-oil cohorts is the baseline. Observations are collapsed at the labor market by phase level. The estimates are weighted by population size. Heteroskedasticity-robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

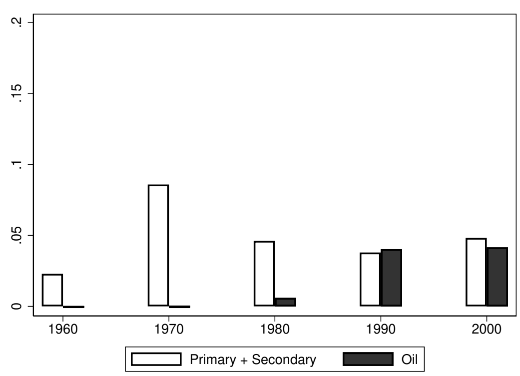
Figure A1: Sectoral Employment and Education for Women, Age 30-35



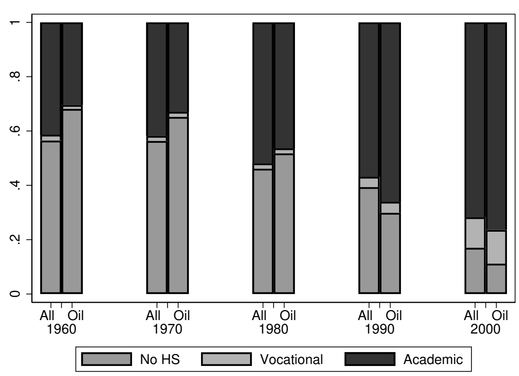
(a) Proportions Employed in Each Sector



(b) Sectors of Employment by Oil Area



(c) Differences in Sectors between Oil Area



(d) Education Levels by Sector of Employment

Notes : Panel (a) shows the proportion of female workers between age 30 and 35 employed in the primary, secondary and tertiary sector (coloured bar, left-axis), and the proportion employed in the oil sector (right axis). Panel (b) shows the same proportions, separately for low-oil ("Low" in the label) and high-oil ("High") regions. Panel (c) shows the difference between high and low oil regions in the proportion of workers employed in the primary and secondary sectors ("Primary+Secondary" in the label) and the difference in the proportion of workers employed in the oil sector ("Oil" in the label). Panel (d) reports the proportion of workers who dropped out of high school ("No HS"), who chose a vocational track, and who chose an academic track, separately for the overall population ("All") and for workers employed in the oil sector ("Oil").

Figure A2: Workers Employed in the Oil Industry in 1980 by Labor Market

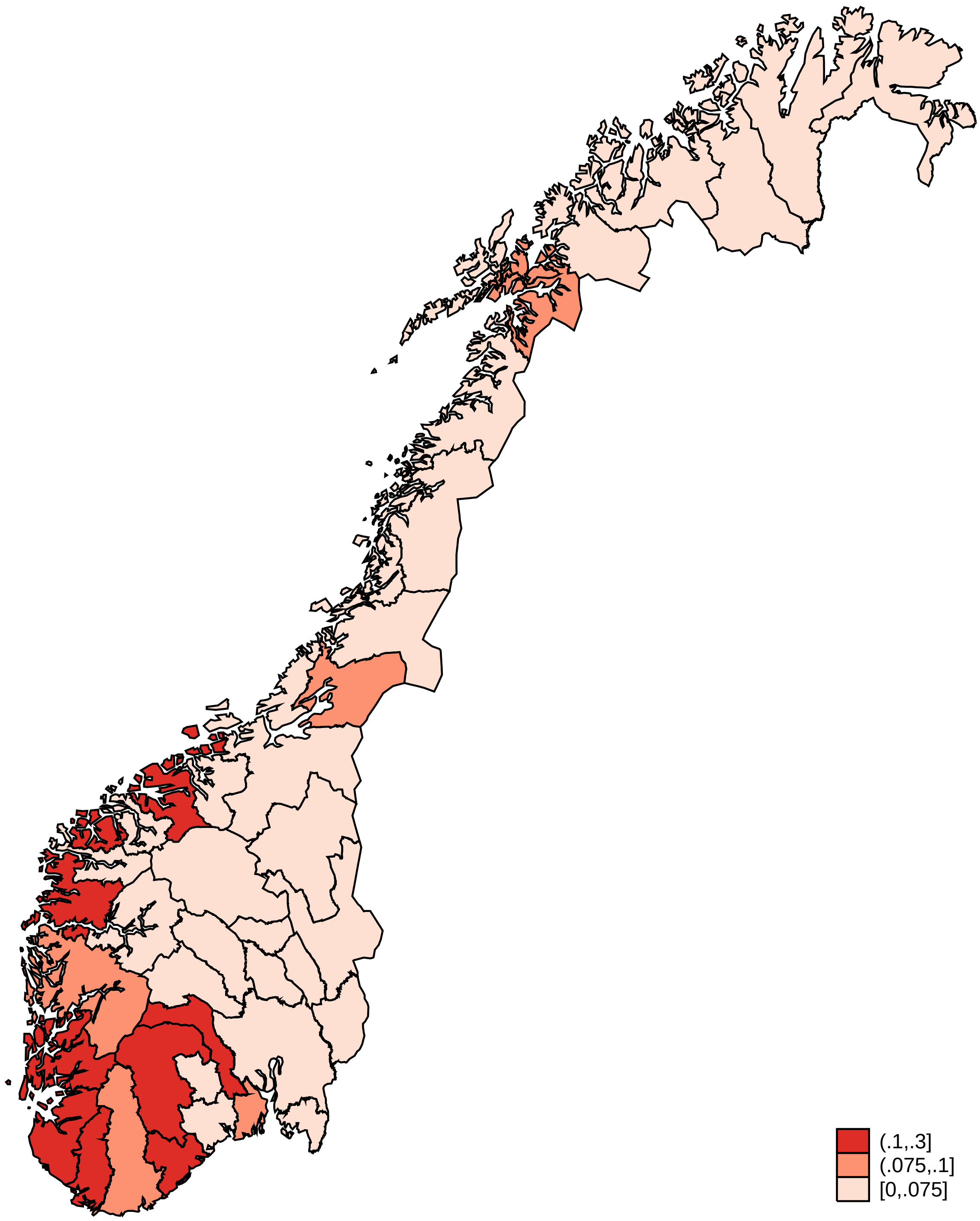
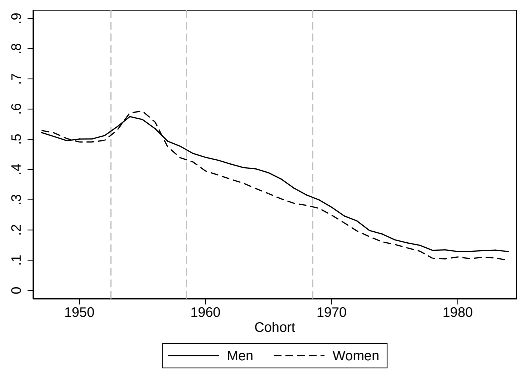
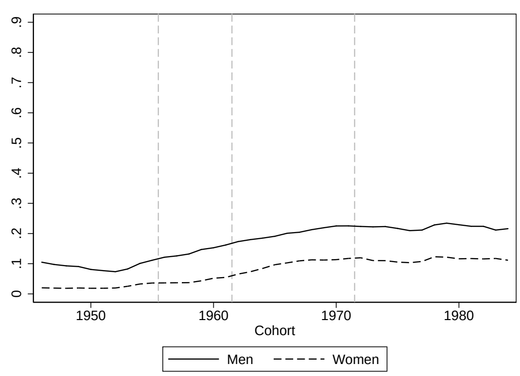


Figure A3: Completion Rates for Different Education Levels By Cohort and Gender



(a) High School Dropout, Age 23

(b) University, Age 30

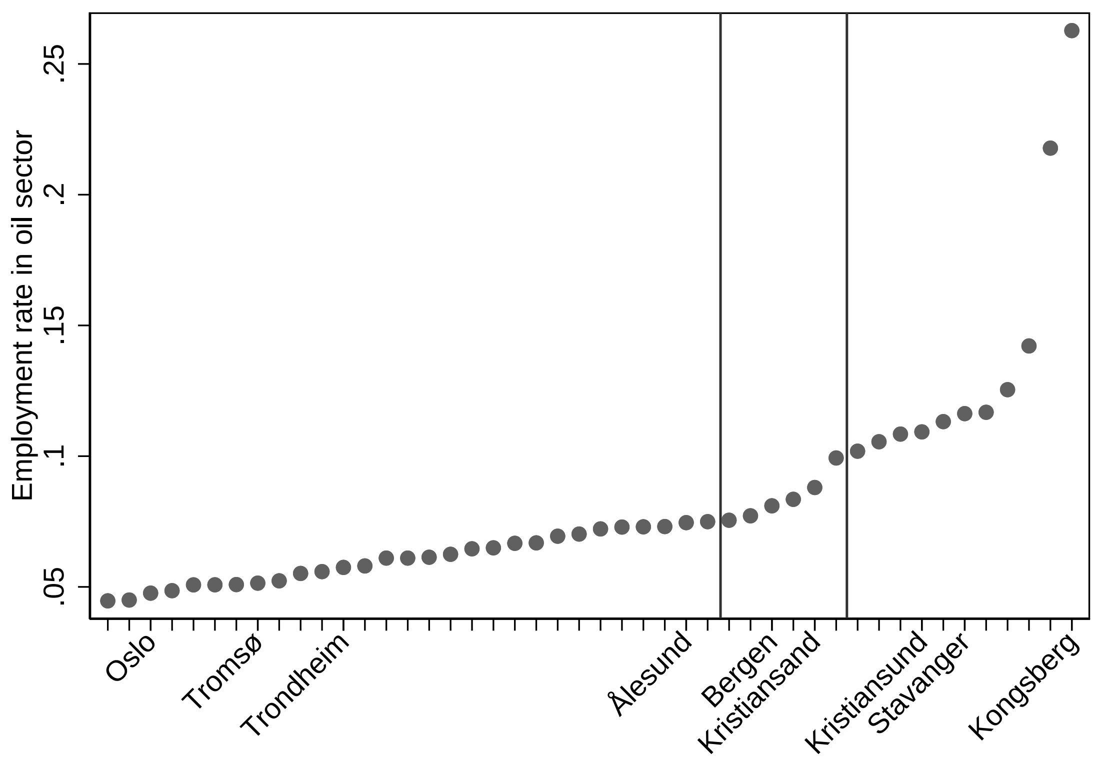


(c) Vocational High School, Age 30

(d) Academic High School, Age 30

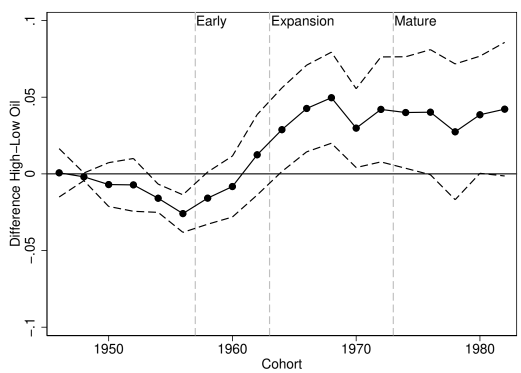
Notes : The figures plot the trend in the raw proportion of students completing different types of education by birth cohort and gender.

Figure A4: Proportion of Oil Industry Workers in 1980 by Labor Market

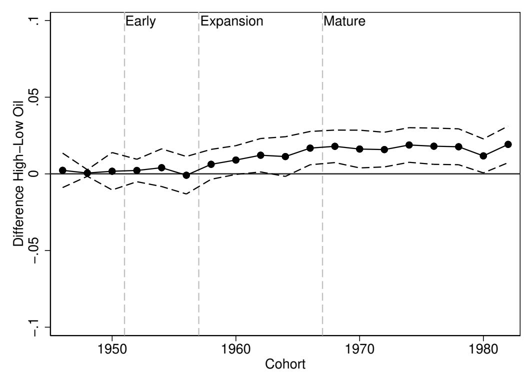


Notes : On the x-axis, labor markets are sorted on the basis of the employment rate in the oil industry. Vertical bars indicate the thresholds that separate low-oil and middle-oil regions (7.5%) and middle-oil and high-oil regions (10%).

Figure A5: Estimated Regional Difference in Vocational High School Completion by Age 23 and Engineering Graduation Rates



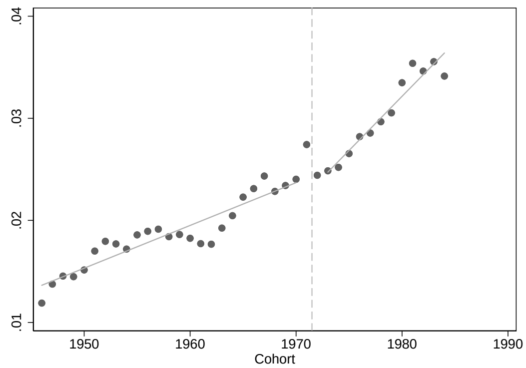
(a) Vocational high school, age 23



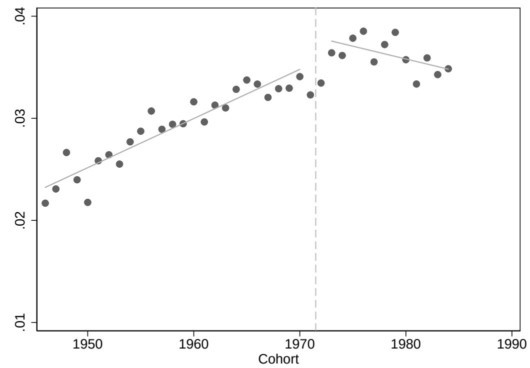
(b) Engineering, whole population

Notes : The figures report the estimates of the parameters for the interaction terms between a dummy for being born in the high-oil region and a dummy for being born in each cohort from Equation (1). The outcome in Panel (a) is a dummy for having completed vocational high school at age 23 and in Panel (b) a dummy for having completed a degree in engineering by age 30 relative to the whole population (and not only relative to those who completed a college degree as in Panel (f) in Figure 2).

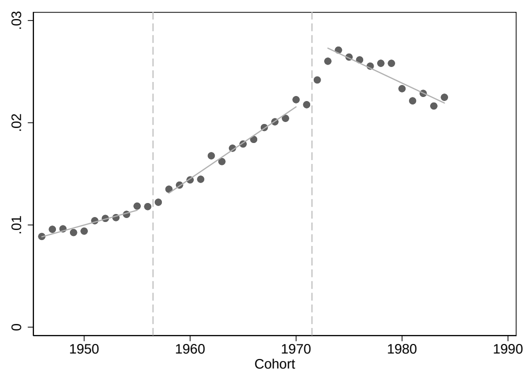
Figure A6: Proportion of Individuals Moving Across Oil Regions



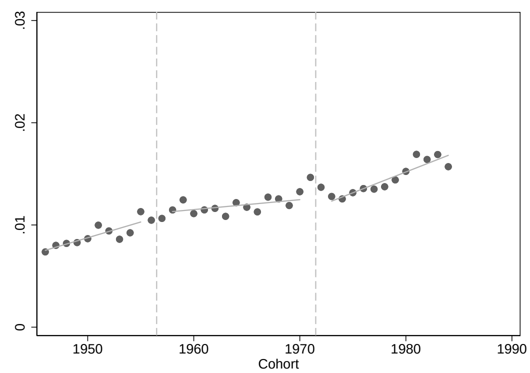
(a) Mothers born in low-oil, children in high-oil



(b) Mothers born in high-oil, children in low-oil



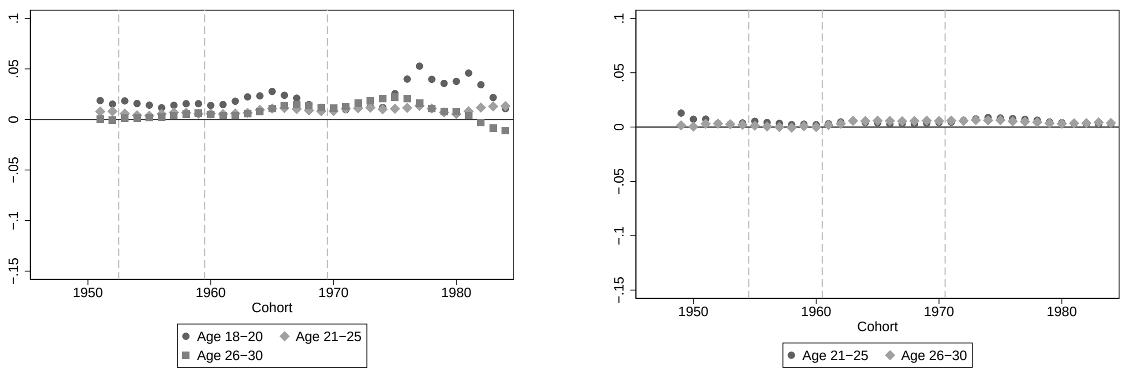
(c) Children born in low-oil, high-oil at age 15



(d) Children born in high-oil, low-oil at age 15

Notes : The figures plot the migration trends between oil areas for individuals born in different cohorts (men and women together). Panel (a) shows the proportion of individuals whose mother was born in a low-oil region who are themselves born in a high-oil region. Panel (b) plots the proportion of individuals whose mother was born in a high-oil region who are themselves born in a low-oil region. These plots proxy the proportion of individuals whose parents migrated before they were born. Panel (c) shows the proportion of children born in low-oil areas who reside in a high-oil area when they are 15 years. Panel (d) plots individuals born in high-oil areas who reside in a low-oil area when they are 15 years. These figures show the proportion of individuals moving while they are children. We observe an increase in individuals moving from low- to high-oil regions for cohorts born after 1955. For cohorts born after 1970, we observe a decline. The number of families migrating is higher before than after birth, suggesting that the relocation decision normally takes place before the birth of a child.

Figure A7: Regional Difference in the Employment Rate by Year, Age, and Education

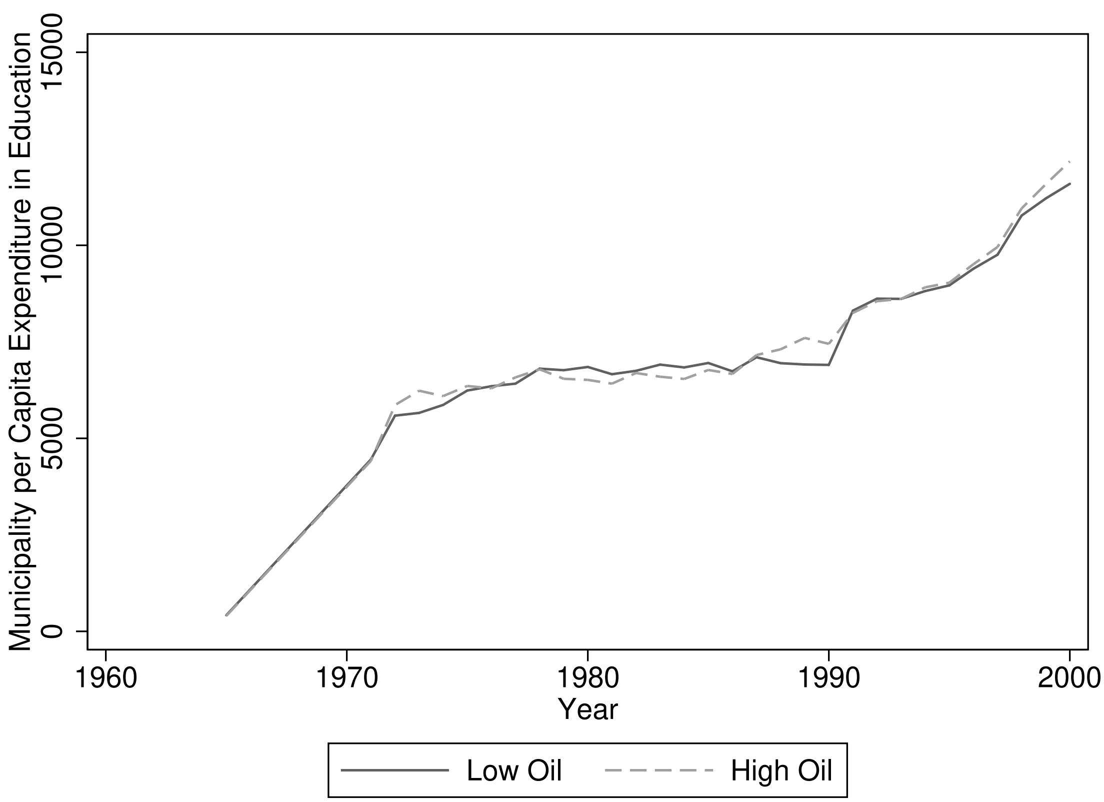


(a) High school dropouts

(b) Vocational high school

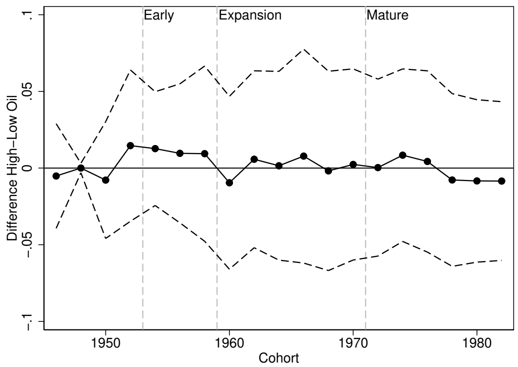
Notes : The figures plot the raw difference in the employment rate (defined as the proportion of men declaring a positive income) between high- and low-oil regions by education levels and age groups. Note that the age group 18 and 20 years are not relevant for vocational high school graduates as they are still in education by that age.

Figure A8: Municipalities' Average Per Capita Expenses on Education by Oil Regions

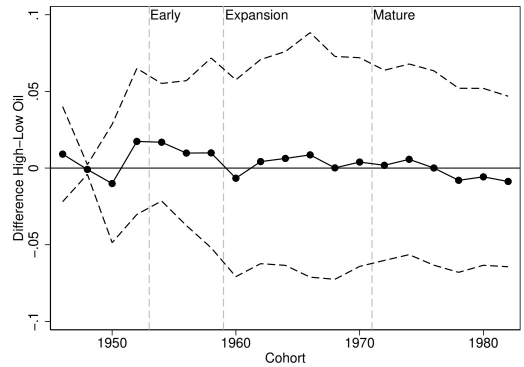


Notes : The data is retrieved
from the municipal accounts and is expressed in constant 1998 NOK. It represents the average amount allocated to education by each municipality in the high- and low-oil region divided by the municipality's population in the year.

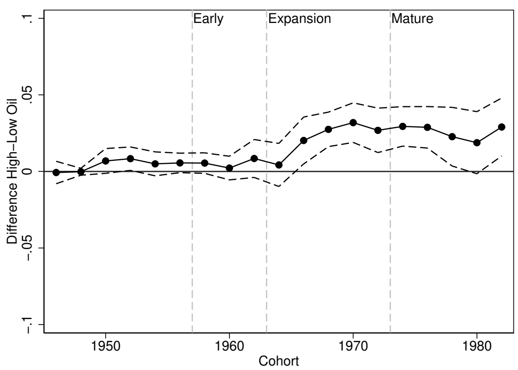
Figure A9: Estimated Difference Between High- and Low-Oil Regions by Educational Level, Women



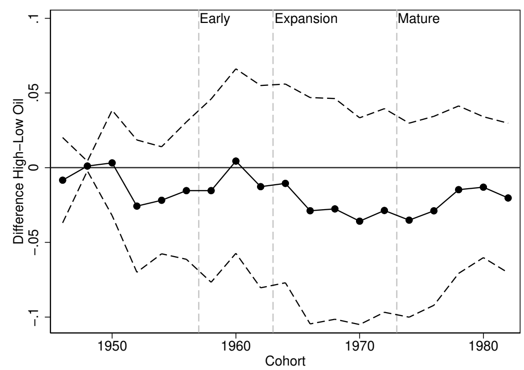
(a) High school dropout, age 23



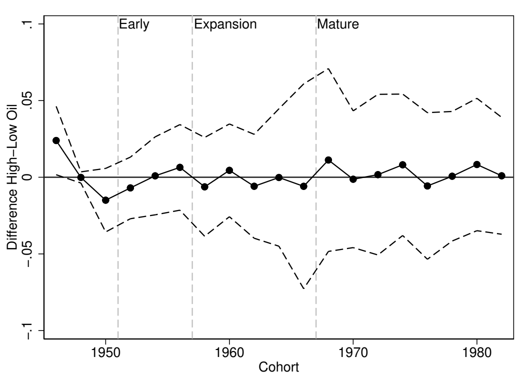
(b) High school dropout, age 30



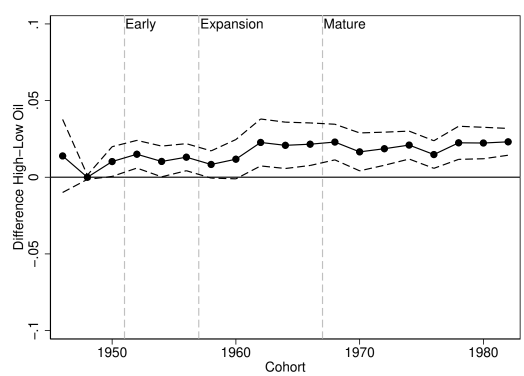
(c) Vocational high school



(d) Academic high school



(e) College



(f) Engineering, college graduates

Notes : The sample includes female students only. The graphs plot the estimates of the parameters from Equation (1) for the interaction terms between a dummy for being born in the high-oil region and a dummy for being born in each birth cohort. The dependent variables are different educational outcomes. Panel (f) includes only students who graduated from college. Where not specified, education levels are measured at the age of 30 years. We exclude the interaction between the high-oil region dummy and the dummy for the pre-oil cohort so that the pre-oil cohorts serve as the baseline. We control for cohort and labor market fixed effect, the population in the municipality of birth in the year the person turns 15 (standardized to mean zero and standard deviation 1 in every cohort), the number of siblings, the birth order, the annual earnings of fathers in the year the individual turns 17, and the highest educational attainment of parents. We cluster the standard errors at the labor market level. Dashed lines show 95% confidence intervals for the estimated parameters. Vertical lines separate the development phase during which the cohorts make different education decisions.

Figure A10: Unemployment Rate and Individuals Reporting Zero Earnings

Notes : The unemployment rate (left-axis) is obtained from the official statistics.
The proportion of individuals who reported zero earnings is taken from administrative tax data (right-axis).