Regional Science and Urban Economics 43 (2013) 164–176

Contents lists available at SciVerse ScienceDirect

Regional Science and Urban Economics

journal homepage: www.elsevier.com/locate/regec

Geographical constraints and educational attainment

Torberg Falch [a], Päivi Lujala [a,b,*], Bjarne Strøm [a]

[a] Department of Economics, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway
[b] Department of Geography, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway

ARTICLE INFO

Article history:
Received 11 January 2012
Received in revised form 26 June 2012
Accepted 26 June 2012
Available online 2 July 2012

JEL classifications:
I21
R23

Keywords:
Dropout
Upper secondary school
Geographical constraint
School location

ABSTRACT

This paper estimates the impact of geographical proximity to upper secondary schools on graduation propensity. It uses detailed information on real travel time between students' homes and schools in Norway and on the composition of study programs at each school. We find that reduced travel time has a positive effect on graduation. The result is robust to a number of specifications, including IV-models and differences-in-differences models. The effect seems to be strongest for students with mediocre prior academic achievement, which suggests that mainly students at the margin of graduation are affected by geographical constraints.

1. Introduction

Policymakers in many countries are concerned with significant proportions of youth dropping out of upper secondary education. A high dropout rate generates inequality, as graduation from upper secondary education is required in most countries for both higher education and certification in several occupations. It can also have detrimental effects on economic growth.

The education level in the population has been found to be important determinant of regional economic growth. Using historical evidence from 19th century Prussia, Becker and Woessmann (2009) show that educational differences causally explain regional income differences, and Becker et al. (2011) find that pre-industrial regional differences in basic education explain a significant part of regional differences in industrialization. de la Fuente and Vives (1995), de la Fuente (2002), and Ciccone and Papaioannou (2009) provide evidence of a substantial impact of education on regional growth in modern economies. Direct evidence on the link between growth and school structure is scarce, but Andersson et al. (2004, 2009) find that decentralization of Swedish higher education increased regional innovation and productivity growth. Thus, geographical constraints on educational choices may be an important determinant of regional differences in education levels and subsequently regional growth.

Within a standard human capital framework, higher costs of attending school reduce educational investment. Individuals living close to universities or upper secondary schools arguably have lower commuting, relocation, or psychological costs. This relationship has motivated several authors to use college or school proximity as instruments for educational attainment in Mincerian earnings equations, following the seminal paper by Card (1995).[1] Recent studies have found that distance to higher education institutions affects participation in higher education and later outcomes.[2]

Completed upper secondary school is a prerequisite for enrolment in higher education. However, the impact of geographical constraints in terms of travel distances between homes and upper secondary schools has received limited attention in the empirical literature.[3] The present paper contributes to the literature by providing evidence of the impact of geographical constraints on the propensity to graduate on time from

* Corresponding author at: Department of Geography, Norwegian University of Science and Technology, 7491 Trondheim, Norway. Tel.: +47 7359 1917; fax: +47 7359 1878.
E-mail addresses: Torberg.Falch@svt.ntnu.no (T. Falch), Paivi.Lujala@svt.ntnu.no (P. Lujala), Bjarne.Strom@svt.ntnu.no (B. Strøm).

[1] Other studies using distance to college as instrument include Kling (2001), Cameron and Taber (2004) and Carneiro et al. (2011). Becker and Siebern-Thomas (2007) use an indicator for growing up in urban areas as instrument in earnings equations for Germany based on the argument that more upper secondary schools are available in urban areas.

[2] See, e.g., Card (1995), Do (2004), Griffith and Rothstein (2009) and Koedel (forthcoming) for the U.S., Sa et al. (2006) for the Netherlands, Frenette (2009) for Canada, and Gibbons and Vignoles (2011) for the U.K.

[3] The only study seems to be Dickerson and McIntosh (2010), who study the impact of distance to upper secondary schools on educational attainment in the U.K.

0166-0462/$ – see front matter
doi:10.1016/j.regsciurbeco.2012.06.007

upper secondary education, using different empirical strategies and detailed Norwegian data on travel time and individual characteristics.

In contrast to most other studies, our measures of geographical constraints are based on travel time between student's residential location and school location, calculated using detailed road map information including data on driving speed limits. Accordingly, our measures should come close to the real geographical constraints facing students. We match this information with data on students' educational careers and individual and family characteristics, and estimate models of on-time graduation from upper secondary education for the cohort leaving compulsory education in 2002. Norway is well suited to the study of geographical constraints on education choices because proximity to schools varies substantially within the country, while at the same time the country is very homogenous in terms of institutions and culture.

Although detailed travel time information and educational data are available, isolating the impact of geographical constraints is challenging because possible unobservable variables might affect both student performance and location of families and schools. We use three approaches to identify the relationship between attainment and geographical constraints. First, we include fixed effects to control for unobservable variables that might be related to labor market opportunities, school district characteristics, and upper secondary school location and quality. Second, we provide instrumental variable estimates using the settlement pattern in the relevant lower secondary school district as an instrument for travel time. While the fixed effect approach identifies the effect of travel time on within-district variation, the IV approach relies on variation across school districts (municipalities). Third, we investigate to what extent the estimated effects are heterogeneous across gender, immigration status, and academic achievement prior to enrolment in upper secondary education. To further study this issue, we use a differences-in-differences strategy to examine whether the effect of travel time is higher for students at the margin of graduation.

The paper is organized as follows: Section 2 reviews relevant literature. Section 3 provides a simple theoretical framework, while Section 4 presents institutional information, the data, and our empirical strategy. The empirical results are provided in Section 5. Section 6 concludes.

2. Literature review

There is a large literature studying the determinants of dropping out of school. An important research question has been to quantify the socioeconomic gradient, i.e., the propensity of students from disadvantaged backgrounds to drop out. Bradley and Lenton (2007) for the U.K., Maani and Kalb (2007) for New Zealand, and Falch and Strøm (2011) for Norway have investigated the role of individual and family characteristics and find that students' prior achievement is a strong predictor of the probability to drop out of non-compulsory secondary education. Similarly, Eckstein and Wolpin (1999) and Belley and Lochner (2007) find that cognitive ability is an important determinant of high school graduation in the U.S. These studies also typically find that students with a disadvantaged background in terms of low parental education and income are more likely to drop out of high school, holding prior achievement constant. Another strand of the literature investigates labor market effects. Rice (1999), Black et al. (2005) and Clark (2011) find evidence from the U.K. and the U.S. that higher regional unemployment and lower relative wages for unskilled workers increase participation in further education after graduating from compulsory school.

Evidence on the impact of geographical constraints on these decisions is, however, very limited. Becker and Siebern-Thomas (2007) find that the supply of high schools is higher in urban areas in Germany and this pattern motivates their use of living in urban vs. rural areas during childhood as an instrumental variable for educational attainment in earnings equations. To our knowledge, the only paper providing a detailed investigation of the effect of distance between home and school on upper secondary education participation is Dickerson and McIntosh (2010), using data from the U.K. On average, the closeness to the nearest school does not affect the decision to participate in full-time upper secondary schooling. The authors find, however, that distance matters for pupils whose grades were mediocre during compulsory education: as the distance to closest school providing academic education increases, these students are less likely to participate in post-compulsory education in general and tend to switch to vocational education. Similarly, they find evidence that distance matters for individuals that have disadvantaged backgrounds.

Our study differs from the Dickerson and McIntosh (2010) study in some important aspects. First, we study graduation from upper secondary education rather than participation one year after the end of compulsory education. Second, our study is based on register data for a complete cohort. Third, we use driving time corrected for speed limits along the road as our distance measure. Dickerson and McIntosh measure distance "as the crow flies", which arguably may differ substantially from real travel distance in the cross-section. If the hypothesis is that close proximity to schools increases educational attainment due to decreasing costs, the relevant measure is the real travel time.

While the evidence on distance effects on upper secondary attainment is limited, there is a literature studying how distance between home and higher education institutions affects education participation and outcomes. Using U.S. data, Card (1995) finds that individuals living closer to four-year colleges attained more years of education than those living further from such institutions. Koedel (forthcoming) finds that states with fractionalized public higher education systems in terms of many small institutions have higher overall university participation, but that they also have a higher exit rate from in-state public institutions to private and out-of-state institutions. The latter finding suggests that a geographically decentralized supply of education may come at the cost of reduced quality.

For the U.S., Do (2004) finds that low-income individuals are more likely to attend a good public college if there is one nearby. Griffith and Rothstein (2009) find similar results using data from the National Longitudinal Survey of Youth 1997. Frenette (2009), using Canadian data, finds that creation of a university in an area increases university participation of local youth, especially among students from lower-income families. For the U.K., Gibbons and Vignoles (2011) find that distance from home to university has a small effect on higher education participation, while distance is the most important factor affecting university choice given participation. The evidence from the Netherlands in Sa et al. (2006) indicates that geographical proximity to universities and professional colleges increases the probability for school leavers to continue education at the post-secondary level. Using detailed geographical data from Sweden, Kjellström and Regnér (1999) find that higher distance between area of residence and nearest university has a small but statistically significant negative effect on university enrolment.

While the existing evidence clearly suggests that distance matters for participation decisions, numerical effects vary substantially between different studies. In addition, there is very limited evidence on the impact of distance and travel time between home and upper secondary schools.

3. Theoretical considerations

Our aim in this paper is to quantify the effect of proximity to schools on the students' propensity to graduate from upper secondary education. A natural point of departure to understand the decisions involved is the standard theory on investment in human capital as originally formulated by Becker (1964) and Ben-Porath (1967). An updated discussion of the theoretical models and empirical work on drop out from upper secondary education is given in Bradley and Lenton (2007). According to the investment theory, a student graduates from upper secondary education if the expected benefit from doing so is higher than the expected cost. The benefit is represented by the expected increase in lifetime income. Costs include the expected forgone earnings when studying, direct costs in terms of tuition, transport and school material, the effort required for graduation, and the risk of failure.

To formalize the mechanisms, assume that student utility is determined by discounted income in the two states, W^G and W^{NG} where superscripts G and NG denotes graduation and non-graduation, respectively, and a set of (choice-independent) individual characteristics, represented by a row vector Z. In addition, there are costs C related to the choices. According to the human capital theory, the student graduates from upper secondary education if the net expected utility of graduation exceeds the alternative. The condition is formally stated for individual i in Eq. (1), normalizing on the cost of the alternative.

$$E\left[U\left(W_i^G, Z_i\right)\right] - C_i^G > E\left[U\left(W_i^{NG}, Z_i\right)\right] \quad (1)$$

This model formulation, which assumes for example perfect information and access to perfect credit markets, implies that a student enrolling upper secondary education also graduates. Thus, the model does not distinguish between enrolment and graduation. However, graduation may differ from initial enrolment if some of the underlying assumptions do not hold. For example, students are likely to be imperfectly informed about their own skills, the effort required to graduate, and the costs and benefits of education. In addition, the job opportunities might change as the students get older. Further, study effort is in practice a choice variable. Students with lower initial skills must expect to put in more effort than students with higher initial skills. These factors may induce dropout and delayed graduation.

With these caveats in mind, we assume that the cost related to graduation is a function of individual characteristics in addition to the travel time between school and home (Q), i.e., $C_i^G = C(Q_i, Z_i)$. Several arguments suggest that the propensity to graduate from upper secondary education is negatively related to travel time. Increased travel time increases money outlays for transportation, and potentially the need to rent a studio. Then the benefit of graduation exceeds the cost for some additional students. In addition, longer travel implies less time for study and potentially more tiredness, which make it harder to exert the effort required to graduate given initial enrolment. Psychological factors such as feelings of distress may add to these cost elements. The suggested negative relationship between graduation and travel time is related to urban efficiency wage theories as laid out in Zenou (2002) building on the assumption that worker effort decreases as commuting distance increases. A recent study by Van Ommeren and Gutiérres-i-Puigarnau (2011) finds that worker productivity, measured by absenteeism, falls when commuting distance between home and work increases.

We will assume that discounted income W^G and W^{NG} in the two states, conditional on individual characteristics Z, varies only across local labor market regions. Framing the model as a reduced form linear probability model, the outcome can be written:

$$y_{ij} = \alpha_j + Q_i\beta + Z_i\gamma + \varepsilon_{ij}, \quad (2)$$

where y_{ij} is an indicator for graduation on-time for individual i in labor market region j, α_j are region fixed effects, β and γ are coefficient vectors to be estimated, and ε is a random error term. While travel time between home and school is our primary variable of interest, we also investigate whether graduation from upper secondary education is affected by broader measures of geographical constraints, such as the number of schools and study tracks within certain travel time thresholds. As our sample is the cohort finishing lower secondary education at the same point in time, our dependent variable combines both the impact of geographical constraints on the students' decision to enroll in upper secondary and the probability that enrolment is successfully converted into graduation.

4. Institutions, data, and empirical strategy

4.1. Institutional background

Compulsory education in Norway consists of seven-year primary and three-year lower secondary education. After finishing lower secondary education, students can either choose to leave education or enroll in one of 15 different study tracks in upper secondary education.[4] The latter alternative is chosen by over 95% of each cohort. After completing the education program in one of these tracks, students get an upper secondary education diploma qualifying for further studies or certifying for work in a number of occupations.

Students enroll in two broad categories of study tracks: academic tracks and vocational tracks. The general academic study track is the largest track and includes about 40% of the total number of enrolled students.[5] The academic study tracks are three-year programs. Vocational study tracks include, e.g., industrial design, health and social work, mechanics, and electrical trades, which certify for work in a number of jobs, such as carpenter and electrician. They are three or four-year programs and most of them include an apprentice system in the third and fourth year, where the training is combined with commercial work in firms.

Municipalities are responsible for compulsory education. The municipalities are multi-purpose authorities and, on average, spending on compulsory education accounts for about 25% of their budgets. Provision of upper secondary education is a county responsibility and is the main service provided by the 19 counties in the country, accounting for over 50% of total county spending. The counties do not locate schools in every municipality. The counties are financed by grants from the central government. Youths have a legal right to enroll in upper secondary education in one out of three individually ranked study tracks, a rule that is followed without exception by each county. Students have a right to five consecutive years of upper secondary education. Students can apply for transfer to another study track after being enrolled. A transfer will delay the student's progress because transfer students usually have to start in the first grade in the new study track.

Students from the same lower secondary school enroll in different upper secondary schools, depending on preferred study track, grades from lower secondary education, and preferences for schools. Most schools offer several study tracks. When the number of applicants exceeds enrolment in a study track, students are ranked based strictly on the grades from lower secondary education.[6] At the end of lower secondary education, the students receive 13 grades in different subjects on a scale from 1 (low) to 6 (high). The average grade varies slightly between subjects, from about 3.5 in mathematics to 4.3 in physical education. In addition, the students have to take a central exit exam in mathematics, Norwegian language, or English language. The overall grade used for ranking the applicants for upper secondary education is the average for all subjects. We use this overall average grade, denoted GPA, as a measure of students' prior achievement in the empirical analysis below.

4.2. Data

The student data are obtained from the National Educational Database in Statistics Norway and consist of all students finishing compulsory

[4] In 2006, the number of study tracks was reduced to 12.

[5] There are two additional minor academic study tracks: "Sports and physical education" and "Music, dance and drama".

[6] In addition, the algorithm takes into account that each student must be enrolled in one of the three study tracks on his or her priority list.

lower secondary education in the spring 2002. The student information includes an identifier for the place of living in 2002 and the upper secondary school in which they enrolled. The information is matched with information about their parents. In the empirical analysis, we use only individuals graduating from lower secondary education at the normal age, i.e., those turning 16 years in 2002 (94.8% of the cohort) and exclude individuals with missing data. Detailed definitions of variables and details on the data reduction are shown in Appendix 1. Table 1 presents descriptive statistics for the sample and variables used in the analyses below.[7]

4.3. Dependent variable

Our outcome variable is a binary variable that equals one if the student has graduated from upper secondary education on time, i.e., within three years after the end of lower secondary education for academic study tracks and within three or four years for vocational study tracks (depending on the actual study track). Table 1 shows that only 56% of the sample graduate on time. Another 10% graduate delayed, but within five years (not shown), while the rest have not graduated five years after leaving lower secondary education. It is of interest to consider these numbers in an international perspective. Cross country comparisons are challenging, however, since upper secondary educational institutions vary greatly across countries. For example, on time graduation from high school education in the US is at age 18, while on time graduation from upper secondary education in Norway is at age 19 or 20. Heckman and LaFontaine (2010) estimate that the graduation rate in the US is about 75% (excluding GED recipients). According to OECD (2011), the successful completion rate in Norway is about the OECD average.

4.4. Measures of geographical constraints

Our primary variable of interest is travel time between home and school. Using detailed data on the public road network that, among other things, identify length and speed limit for each road segment, we calculate the travel time by car from home residence to the closest school. In addition, we calculate the number of schools and different study tracks within several travel time thresholds from the student's home residence.

Students' residence is registered at the ward level at January 1, 2002. As Norway is divided into about 14,000 wards ("Grunnkretser"), this gives a reasonably good localization of the students' residence at the end of compulsory education. On average there are about 350 inhabitants in each ward. The sample has 10,857 wards with student observations. The number of students in the relevant cohort in each ward varies from zero to 54. At the ward level, the median is three students and the mean is 4.7 students, while at the student level the median and mean are seven and 9.8, respectively. To make the home residence location even more accurate, we use ArcGIS to determine the midpoint of the ward's populated area. This enables us to exclude all uninhabited parts of the wards, and ensure that the ward midpoint is located near a public road. Fig. 1 illustrates the approach: the light violet areas are the inhabited areas for which the ward midpoint (small dots) is calculated. For the school location (large dots) we use geographic coordinates.

Using ArcGIS Network Analyst, we calculate the travel time between students' home ward midpoint and the nearest upper secondary school. In contrast to the Dickerson and McIntosh (2010) study, which uses distance "as the crow flies", we measure the distance over the public road network. This gives us a more realistic picture of the true travel distance. Moreover, we add precision by using travel time defined as driving time by car at the speed limits rather than the absolute distance. In Fig. 1, the brown straight lines connect the ward midpoint and the nearest school to which there is a road connection.[8]

Fig. 2 illustrates the variation in access to school in the middle part of Norway. Students in wards indicated with red have over 60 min travel time to the closest school,[9] or need to take a boat or ferry. Those living in wards indicated with orange have a travel time between 30 and 60 min. Geographically, red and orange areas cover extensive areas; these areas usually are sparsely populated. Students in more densely populated areas typically have shorter travel distance to the closest school (green and yellow areas on the map).

The travel time to the nearest school varies from close to 0 to 3 h. Table 1 shows that the average travel time to the nearest school is 0.17 h. The distribution is presented in the upper part of Fig. 3. Few students have a long trip to the nearest school because the majority lives in cities; 70% have less than 10 min travel time and 93% have less than 30 min travel time. To illustrate the distribution for the

Table 1
Summary statistics.

	Mean	Standard deviation
On-time graduation	0.56	–
Travel time to nearest upper secondary school, hours	0.17	0.22
Number of different study tracks within 0.5 h travel time	11.3	4.50
Number of upper secondary schools within 0.5 h travel time	14.1	16.8
No upper secondary school within 0.5 h travel time	0.07	–
1-5 upper secondary schools within 0.5 h travel time	0.36	–
No road to upper secondary school	0.01	–
GPA	3.89	0.84
Girl	0.49	–
First generation immigrant	0.04	–
Second generation immigrant	0.02	–
Birth month	6.41	3.35
Parents married	0.63	
Parents divorced	0.12	
Parents never married	0.25	
None of the parents have upper secondary education	0.15	–
At least one parent upper secondary education	0.47	–
At least one parent bachelor degree	0.28	–
At least one parent master degree	0.10	–
Parental income in quartile 1	0.25	
Parental income in quartile 2	0.25	
Parental income in quartile 3	0.25	
Parental income in quartile 4	0.25	
Student moved between municipalities at age between 6 and 16	0.11	
Student mobility unknown	0.02	
Benefits due to disabilities before age 18	0.02	–
Benefits due to disease before age 18	0.02	–
Urban ward	0.74	
Rural ward	0.25	
Urban/rural ward unknown	0.004	
Number of study tracks at nearest upper secondary school	4.67	2.71
Log (Number of students in same cohort in the ward)	1.94	0.86
Average GPA in the ward	3.89	0.44
Share in the ward with at least one parent more than compulsory education	0.85	0.18
Share immigrants in the ward	0.05	0.13
Square km per inhabitant in the municipality	0.08	0.20
Share of population living in rural areas in the municipality	0.25	0.24
Observations	51,484	

[7] 97.0% of the sample used in the empirical analysis enroll in upper secondary education the year they finish compulsory education. For the population of individuals finishing compulsory education in 2002, that is the case for 95.1%.

[8] Wards that have no road connection to a school are coded separately in the analyses below. These are almost exclusively islands without schools which are not connected to the mainland by bridge or tunnel, and include 1.2% of the students in the sample. One such island is in the south-west corner of Fig. 1.

[9] Only 1.3% of the students in the sample have road connection to a school but more than 1 h travel.

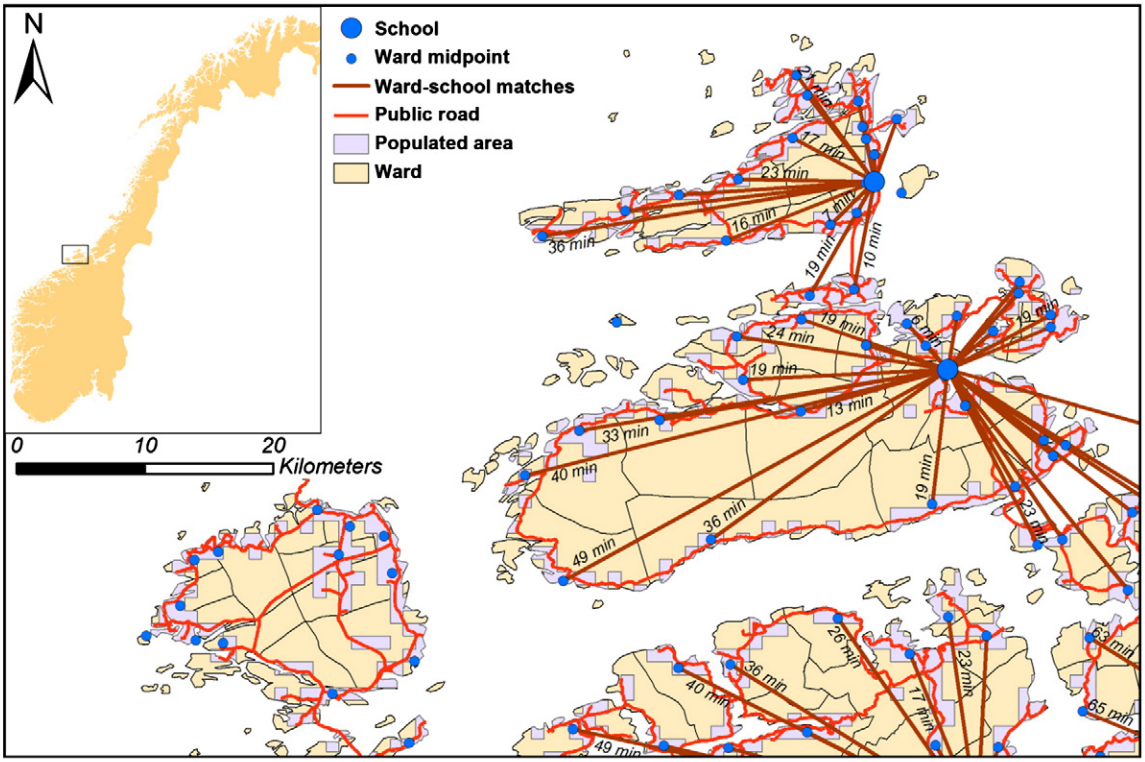

Fig. 1. Closest school and travel time by car using public roads and following speed limits in a coastal area of middle Norway. The brown lines connecting ward midpoints (small dots) and schools (large dots) indicate the closest school to each ward. The island on the left is not connected to the mainland by bridge or tunnel and the relevant wards are thus not assigned a travel time.

majority of the students more clearly, the upper right part of the figure shows the distribution for students with less than 30 min travel time.

Naturally, the feasibility of driving at the speed limit depends on the number of crossings, traffic lights, etc. Since the speed limit typically is low on roads with many obstacles, our measure is still a substantial improvement of the Euclidian distance.[10] In practice, students use different transport modes. Actual travel time will, of course, differ from our measure if the student uses a bicycle or walks to school, which they typically do when they live nearby their school. With long distances it is typically possible to use public transport, but it is also common that parents drive their children to school. In addition, many students obtain the driving license during the period in high school education. The actual choice of transport mode presumably varies across neighbors and depends on a wide range of family conditions in addition to distance to school, and might vary over the year due to weather conditions. In addition, the student might not be enrolled at the school closest to their residency as discussed above. Our travel time variable can thus be interpreted as an objective measure of the minimum travel time for a representative student in each ward.

In addition to travel time between home and nearest school, we create a list of schools that are within certain driving distance thresholds from students' home ward midpoint. In the analysis below, we will mainly rely on the 30 min travel time threshold. In this case, the distance analysis yields over 145,000 matches between ward midpoints and schools. Based on travel time thresholds, we create two broader measures of geographical constraints. The first is simply the *number of schools* that fall inside the travel time definition. The second is the *number of unique study tracks* offered in the upper secondary schools located within the travel time definition.[11] This reflects the study program choice set for the student, and varies from zero to 15. This variable combines information on proximity to schools, the supply of study tracks, and the composition of the study tracks.

Using the 30 min travel time constraint, Table 1 shows that students are faced with a supply of 11.3 different study tracks on average. The cities have typically a complete supply of study tracks. The number of schools within the 30 min travel time constraint varies from zero to 59, with the largest number of schools available in the Oslo area (the capital). Table 1 shows that a random student has 14.1 upper secondary schools within 30 min travel. Seven percent of the students have no schools in the vicinity, while 36% have 1–5 schools within 30 min travel time.

4.5. Individual and family control variables

We include the average grade in compulsory education as a measure of prior achievement. This is an important variable because it controls for the initial sorting of students and families across geographical areas. In addition, we include indicators for student's gender, immigration status, birth month, mobility across municipalities during compulsory education, the need for support related to disabilities before age 18, residency in urban vs. rural ward at age 16; civil status, highest education, and income quartile for the parents when the student is 16 years old; and some neighborhood peer variables

[10] The correlation between our travel time measure and the Euclidean distance is 0.86 (excluding the observations with missing travel time information). The correlation is mainly driven by the densely populated areas. The correlation coefficients for observations with more than 15 min and more than 30 min travel time are 0.66 and 0.44, respectively.

[11] That is, if two or more schools provide the same study track, the study track is counted only once.

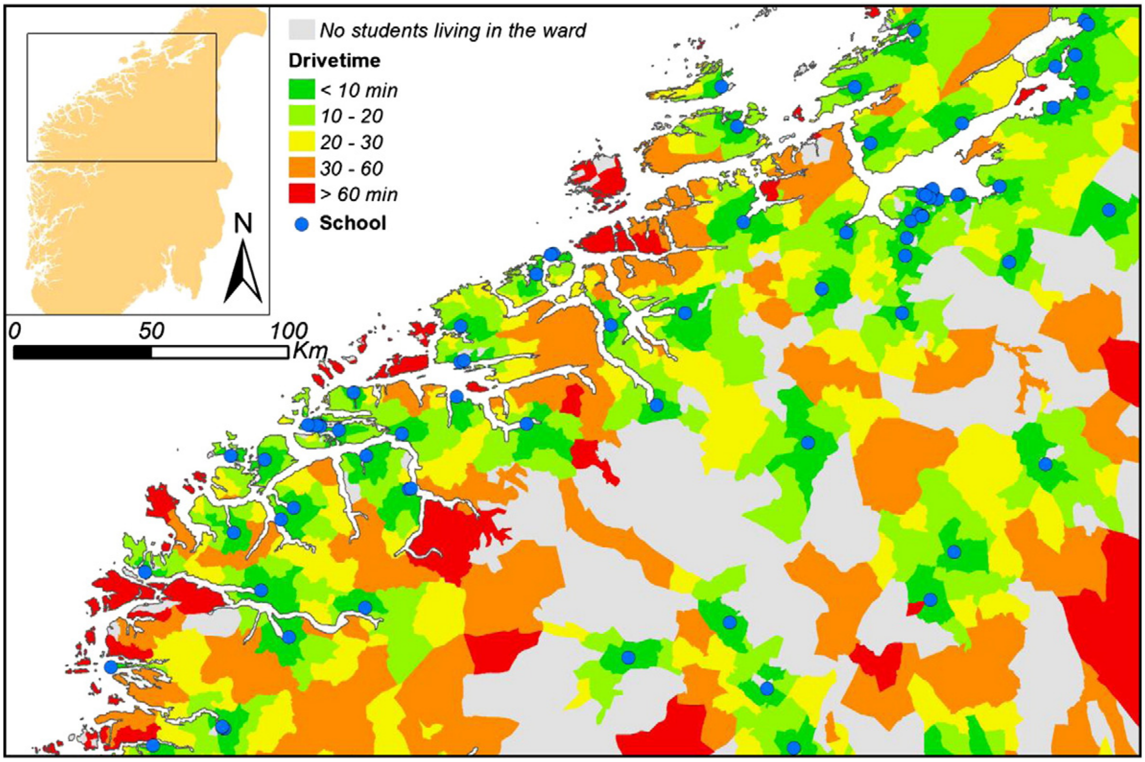

Fig. 2. Travel time from the ward midpoint to closest school in middle Norway. Areas indicated with the grey color are uninhabited or did not have any students leaving the lower secondary school in 2002.

measured at the ward level and, there are 4% first generation immigrants and 2% second generation immigrants in the sample. 63% of the parents are married, 12% are divorced and 25% have not been married. For 15% of the students, neither of the parents has education above the compulsory level, while, for 10%, at least one of the parents has a master degree.

4.6. Empirical strategies

Our empirical task is to isolate the impact of geographical proximity to school on student's probability to graduate on time. This is challenging because unobservable variables might affect both student performance and the location of families and schools. Variables

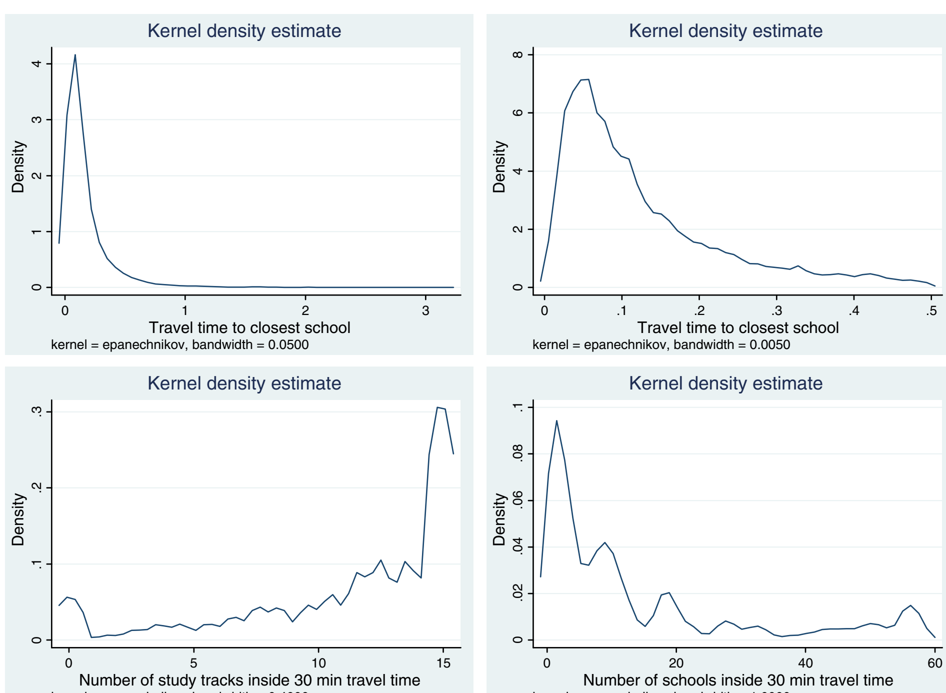

Fig. 3. Density of measures for geographical constraints: Driving time to nearest school for the whole sample and for those with travel time of less than 30 min (upper figures); number of unique study tracks within 30 min travel time; and number of schools within 30 min travel time (lower figures).

characterizing school structure vary little over time, limiting the scope for identification by use of conventional panel data methods or policy interventions. Our identification relies on variation in the cross-section dimension, and we explore different empirical strategies.

At the outset we restrict the analysis to the students' choice set in the proximity of their residence at the end of compulsory education. In this sense we estimate a reduced form model. Both residential mobility in terms of students living outside their parental homes while enrolled in upper secondary education and the actual school in which the students enroll can clearly be endogenous. Actual enrollment depends on individual preferences and GPA from compulsory education. Enrolment at the neighboring school might not be an option if that particular school is oversubscribed. Most important, it is not uncommon to freely choose another school than the neighboring school. Since we do not use the travel time between the students' home and their actual upper secondary school, we also extend the analysis to include broader travel time measures, i.e., the number of schools and the number of unique study tracks within certain travel time thresholds.

We initially estimate the baseline model described in Eq. (2) above, including a host of individual variables along with the travel time variable. The omitted variable problem is dealt with by three different approaches; fixed effects, IV-estimation, and a specific differences-in-differences specification.

We estimate two fixed effects models. Including fixed effects for the 433 municipalities in which the students lived at the end of compulsory education captures a substantial part of the variation in unmeasured geographical characteristics. Several municipalities have no upper secondary schools located within their boundaries. Second, we expand the model by including fixed effects for each of the 501 upper secondary schools in the sample, which effectively control for unmeasured school quality related to graduation in addition to some geographical factors.

Our second approach is to use IV-estimation. The IV-approach requires variables that affect proximity to schools and study tracks, while having no direct effect on the probability of graduating, conditional on the included control variables. Our approach is to use municipal characteristics as instruments. This is based on the argument that the number of upper secondary schools in a district depends on the settlement pattern of the municipality. The county is unlikely to establish upper secondary schools in municipalities with scattered population. We use the share of the population living in rural areas and square kilometers per inhabitant in the municipality as instruments for the travel time from home to the closest upper secondary school.[12] Table 1 shows that there is substantial variation in these variables.

Consistency of the IV-estimator requires that the proposed instruments have a significant impact on distance to school (instrumental relevance), while there is no direct impact on graduation on time, conditional on the included variables (exclusion restriction). One possible concern with our strategy is that population density in the municipality of the student's lower secondary school might itself be correlated with factors affecting student performance. For example, students from low-income families may settle in certain areas in the municipality. Including the ward level average student achievement, parental income, and parental education in the model should, however, reduce this problem.

These two approaches use different variation to identify the effect of travel time. While the fixed effects approach rely purely on within-municipality variation, the IV-approach exploits travel time variation determined by variables at the municipality level for identification.

The third approach examines heterogeneous effects of travel time. Students in the upper part of the ability distribution are likely to graduate anyway, while students in the lower part of the ability distribution are highly likely to drop out independent of locational characteristics. The fact that on average 56% graduate on time suggests that it is students with mediocre skills that are on the margin of graduation. Additional costs of schooling related to geographical constraints are arguably most important for these students.[13] We find such a pattern, and explore this finding using fixed effect strategy. We estimate a model with fixed ward effects along with interaction terms between travel time and indicators of the students' quartile in the ability distribution. This can be interpreted as a differences-in-differences strategy in order to investigate whether the effect of travel time is higher for students at the margin of graduation than for other students. This approach does not identify the general effect of travel time, but it differences out all geographical factors to the extent that the sorting behavior is similar across the ability distribution.

5. Empirical analysis

5.1. Fixed effects approach

We begin by estimating variants of Eq. (2) by OLS using travel time to the nearest school as a measure of geographical constraint. Results are presented in Table 2, where the dependent variable is the probability of on time graduation. For comparison purposes, column (1) shows the result when travel time is the only explanatory variable in the model, while column (2) adds 90 regional dummy variables to capture the effect of external labor market conditions. The regions are defined by Statistics Norway using commuting statistics. The regional fixed effects are likely to capture both differences in labor markets and the fact that geography varies across the country. Both specifications show a significant negative correlation.[14]

Column (3) adds a number of individual characteristics to the model. The effects of these variables are mainly as expected. GPA from lower secondary education is a very strong predictor of graduation. Increasing GPA by one unit (on a scale from 1 to 6) increases the probability of graduating on-time by 31 percentage points.[15] Girls and immigrants have a higher probability of graduating, and there is a positive effect of birth month conditional on GPA. The latter must be interpreted as a catching up effect since there is a negative effect of birth month on GPA. Married, well-educated and rich parents increase the graduation probability, while mobility has a negative effect. In addition, graduation decreases in average GPA in the ward, the share of immigrants in the ward, and the share of low-educated parents in the ward.

While the covariates add significant explanatory power to the model, the effect of proximity to school is only moderately reduced in numerical terms when the individual characteristics are included in the model. The estimated coefficient suggests that students with 30 min travel to the nearest school have 2.3 percentage points lower probability of graduating on-time than students with a school across the street.

Column (4) in Table 2 includes municipal fixed effects for the municipality in which the student graduated from lower secondary

[12] Notice that the municipalities typically have several lower secondary schools. On average there are 2.6 lower secondary schools per municipality. Thus, it is not the characteristics of the catchment area of each lower secondary school that are used as instrument.

[13] In contrast, if the graduation rate on average was 90%, the students on the margin of graduation would most likely be at the lower end of the ability distribution.

[14] Logit and probit models are straightforward alternative estimation methods to the linear probability model. However, using those methods, the marginal effects and standard errors at mean values are identical to that reported in columns (1) and (2) in Table 2.

[15] Since the effect of GPA is very strong (t-value of about 90), one may wonder whether the effect of travel time is sensitive to the linear functional form presented. We have estimated several alternative model formulations, including a cubic in GPA and dummy variable specifications, which give nearly exactly the same effect of travel time.

Table 2
Baseline model for on-time graduation.

	(1)	(2)	(3)	(4)	(5)
Travel time to nearest upper secondary school, hours	−0.087*	−0.056*	−0.046*	−0.045*	−0.037*
	(−4.48)	(−3.25)	(−3.65)	(−2.54)	(−2.04)
GPA	–	–	0.31*	0.31*	0.28*
			(89.5)	(89.2)	(82.6)
Girl	–	–	0.021*	0.021*	0.026*
			(3.34)	(3.28)	(4.26)
First generation immigrant	–	–	0.033*	0.033*	0.023
			(2.14)	(2.12)	(1.52)
Second generation immigrant	–	–	0.041*	0.043*	0.040*
			(3.58)	(3.74)	(4.03)
Birth month	–	–	0.0041*	0.0042*	0.0039*
			(8.25)	(8.27)	(7.74)
Parents married	–	–	0.062*	0.061*	0.058*
			(12.9)	(12.5)	(12.4)
Parents divorced	–	–	0.016*	0.015*	0.012*
			(2.76)	(2.60)	(2.04)
At least one parent upper secondary education	–	–	0.022*	0.023*	0.021*
			(3.27)	(3.37)	(3.31)
At least one parent bachelor degree	–	–	0.035*	0.034*	0.031*
			(5.46)	(5.26)	(4.68)
At least one parent master degree	–	–	0.033*	0.031*	0.026*
			(3.30)	(3.09)	(2.51)
Parental income in quartile 2	–	–	0.017*	0.017*	0.013*
			(2.90)	(2.97)	(2.19)
Parental income in quartile 3	–	–	0.031*	0.031*	0.025*
			(5.55)	(5.68)	(4.47)
Parental income in quartile 4	–	–	0.038*	0.038*	0.028*
			(6.85)	(6.96)	(4.67)
Student moved between municipalities at age between 6 and 16	–	–	−0.037*	−0.036*	−0.033*
			(−5.42)	(−5.16)	(−4.92)
Student mobility unknown	–	–	−0.022	−0.022	−0.0055
			(−1.65)	(−1.68)	(−0.38)
Benefits due to disabilities before age 18	–	–	0.0086	0.010	0.0095
			(0.57)	(0.67)	(0.60)
Benefits due to disease before age 18	–	–	−0.062*	−0.061*	−0.059*
			(−4.07)	(−3.96)	(−3.79)
Rural ward	–	–	−0.0074	−0.0057	−0.0014
			(−1.42)	(−0.98)	(−0.24)
Urban/rural ward unknown	–	–	0.024	0.035	0.039
			(0.72)	(1.03)	(1.17)
Log (Number of students in same cohort in the same ward)	–	–	0.0035	0.0038	0.0033
			(1.34)	(1.34)	(1.10)
Average GPA in the ward	–	–	−0.0091*	−0.0047	−0.0077
			(−2.21)	(−1.13)	(−1.85)
Average share of parents with at least upper secondary education in the ward	–	–	0.030*	0.026	0.019
			(2.28)	(1.89)	(1.49)
Average share of immigrants in the ward	–	–	−0.035*	−0.036*	−0.016
			(−2.21)	(−2.37)	(−0.79)
Region fixed effects	No	Yes	Yes	–	–
Municipality FE	No	No	No	Yes	Yes
Upper secondary school FE	No	No	No	No	Yes
Observations	51,484	51,484	51,484	51,484	51,484
R-squared	0.002	0.016	0.324	0.331	0.354

Note: t-values based on standard errors clustered at the regional level in parentheses. Region and municipality fixed effects are assigned according to residence at January 1 in 2002, while upper secondary school fixed effects are assigned to the school in which the individual enrolled in the fall in 2002. The individuals that did not enroll upper secondary education in 2002 are assigned a separate identifier.

* Denotes statistical significance at 5% level.

education.[16] Since the municipalities are responsible for compulsory education, these fixed effects capture important geographical differences during adolescence. Finally, column (5) adds fixed effects for upper secondary schools, effectively accounting for the impact of school quality and school resources. In this model, only within-municipal variation in proximity to the nearest school for students attending the same upper secondary school is used for identification of the impact of travel time.[17] While the coefficient of interest hardly changes when municipal fixed effects are included, it is reduced somewhat when upper secondary school fixed effects are included (from −0.046 to −0.037).

[16] Notice that the municipal fixed effects absorb the regional fixed effects since the municipalities are a smaller geographical unit than the regions.

Fig. 4. Non-parametric model for travel time to nearest school in minutes. Estimates ±2 standard errors and the estimated linear effect (the straight line).

Fig. 4 presents non-parametric estimates of the effect of travel time using the model specification in column (3) in Table 2. The continuous travel time variable is replaced by dummy variables for each 5 min interval between 0 and 60 min, and dummy variables for each 15 (30) minute interval above 60 (90) minutes. Neighborhoods that have to use boat to the nearest school are included as a separate category. The semi-parametric approach more generally addresses the concern that travel time measured by use of car driving at the speed limit might be a poor proxy for the travel modes actually chosen. The straight line in the figure is the estimated linear effect.[18] The figure seems to support a linear representation. However, a formal test adding the dummy variables to the linear model gives a p-value of 0.001. Nevertheless, it is reassuring that the figure indicates that the effect of travel time to the nearest school is not driven by outliers with extremely long travel distances, because the impact using the dummy variable approach is very close to linear in the 0–40 min travel time interval.[19] This evidence indicates that our travel time variable based on car use, which is the minimum travel time, is a reasonably proxy of actual travel time.

5.2. Instrumental variable approach

The most comprehensive models above use within-municipality variation to identify the effect of travel time. However, it is possible that the estimated effects represent omitted characteristics of students and neighborhoods correlated with the measures of geographical constraints. In this section, we therefore estimate the model using

[17] Notice that the students do not need to attend the nearest school. Most importantly, school choice depends on choice of study track, and no school offers all the 15 different tracks. In fact, only six schools in the sample have more than 10 different study tracks.

[18] The line representing the linear effect estimated in column (3) in Table 2 is included in the figure only for travel time up to 60 min. In this interval, each dummy variable represents a travel time change of 5 min. For longer travel time, each dummy variable represents a larger change in travel time, because there are relatively few observations within each group.

[19] Fig. 4 indicates that a logarithmic specification of the travel time variable might yield a better fit to the data. We have estimated the model in column (3) in Table 2 adding the log of travel time. Then neither of the two travel time variables (the linear and the semi-log-linear effects) is significant at 10% level, which implies that the information in the data cannot be used to distinguish between these two functional forms of the relationship.

Table 3
Instrumental variable estimates.

	(1) Reduced form on-time graduation	(2) Reduced form travel time (1. stage)	(3) Structural model on-time graduation	(4) Reduced form on-time graduation	(5) Reduced form travel time (1. stage)	(6) Structural model on-time grad.
Travel time to nearest upper secondary school, hours	–	–	−0.067* (−2.30)	–	–	−0.061* (−2.56)
Square km per inhabitant in municipality	−0.021 (−0.94)	0.25* (2.18)	–	−0.033* (−2.28)	0.35* (3.36)	–
Share of population in municipality living in rural areas	−0.015 (−1.13)	0.29* (5.53)	–	−0.010 (−0.75)	0.32* (6.54)	–
Upper secondary school FE	No	No	No	Yes	Yes	Yes
R-squared	0.323	0.871	0.313	0.349	0.887	0.349
F-test for weak instruments	–	42.3	–	–	61.9	–
Test for overidentifying restrictions (p-value)	–	–	0.718	–	–	0.306

Note: Same model specification as in column (3) in Table 2 except as indicated. 51,484 observations. t-values based on standard errors clustered at the regional level in parentheses.
* Denotes statistical significance at 5% level.

the instrumental variable approach. Our instrument set effectively identifies the coefficient of interest by exploiting the part of variation in student travel time explained by variables that vary only across municipalities. Thus, our IV-approach clearly utilizes a different source of variation to identify travel time effects than the municipal fixed effects specification above.

Based on the arguments in Section 4, settlement pattern in the municipality is used for identification. The instruments are the share of population living in urban areas and the square kilometer per inhabitant. Table 3 reports estimated effects for the variables of interest, while complete results for the model including school fixed effects are provided in Appendix 2.

The first part of Table 3 presents IV results for the model with regional fixed effects and the same control variables as in the OLS model. Column (1) shows the estimated reduced form equation for on-time graduation, column (2) presents the first stage (reduced form) equation for travel time, while column (3) reports the IV-estimates for the structural model. The first stage results show that our instruments have a relatively strong effect on travel time (F-value of 42.3), indicating that the instruments are relevant. The structural estimate in column (3) is significant at the 5% level, and slightly larger than the OLS result (−0.067 vs. −0.046). To provide statistical evidence on the credibility of the exclusion restrictions, we perform formal overidentification tests. The test statistic on overidentifying restrictions is well below critical values and we cannot formally reject the validity of the instruments. However, when interpreting these results, one should keep in mind that the power of overidentification tests is in general low, in particular when the instrumental variables to some extent measure the same underlying factors. In our case, the correlation coefficient between the instruments is 0.46, but both variables are significant in the first stage.

Columns (4)–(6) in Table 3 present the corresponding results for the school fixed effects model. In this case, identification exploits variation in travel time for students who attend the same upper secondary school but live in municipalities with different settlement patterns. Thus, it is mainly variation across neighboring municipalities that is utilized for identification. Again, the first stage results suggest that the instruments are highly relevant, and the overidentification test suggests that the exclusion restrictions are valid. The structural model estimate in Column (6) implies that half an hour increase in travel time to the nearest school increases the probability of graduating on time by about three percentage points, which is a non-negligible effect.

5.3. Alternative measures of geographical constraints

While the estimations so far suggest that geographical constraints in terms of travel time between home and nearest upper secondary school have an impact on educational choices, a relevant issue is to disentangle the separate impact of alternative measures of such constraints.

Table 4 presents results from some models using alternative measures of geographical constraints. As expected given the evidence

Table 4
Different measures of geographical constraints. Dependent variable is on-time graduation.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Travel time to nearest upper secondary school, hours	–	–	–	–	−0.037* (−2.57)	−0.038* (−2.57)	−0.041* (−2.27)	−0.027 (−1.29)	−0.026 (−1.28)
Number of different study tracks within 30 min travel time	0.0028* (3.87)	–	0.0027* (2.01)	–	0.0012 (1.54)	0.0010 (1.00)	0.0014 (1.32)	0.0014 (0.90)	0.0314 (0.82)
No of different upper sec. schools within 30 min travel time	–	0.0006* (3.52)	–	0.0006 (0.59)	–	0.0002 (0.72)	–	–	0.0001 (0.09)
Dummy for no upper sec. schools within 30 min travel time	–	–	–	–	–	–	0.007 (0.33)	–	–
Dummy for 1–5 upper sec. schools within 30 min travel time	–	–	–	–	–	–	0.0001 (0.01)	–	–
Municipality FE	No	No	Yes	Yes	No	No	No	Yes	Yes
Upper Secondary school FE	No	No	Yes	Yes	No	No	No	Yes	Yes
R-squared	0.324	0.323	0.354	0.354	0.324	0.324	0.324	0.354	0.354

Note: Same model specification as in column (3) in Table 2 except as indicated. 51,484 observations. t-values based on standard errors clustered at the regional level in parentheses.
* Denotes statistical significance at 5% level.

above, column (1) shows that the number of unique study tracks within 30 min travel time has a positive and significant effect on graduation, and column (2) shows that the number of schools within 30 min travel time has a positive and significant effect. The former result implies that increasing the number of unique study tracks by 10 increases the probability of graduating by 2.8 percentage points. Including municipal and school fixed effects in columns (3) and (4) do not change the estimates, but makes the effect of the number of schools less precise imprecise.

The three different measures of geographical constraints used so far are correlated. The additional models in Table 4 show that the effect of the number of schools is dominated by the two other measures of geographical constraints. That is the case also when the number of schools is replaced by dummy variables for no schools and 1–5 schools within 30 min travel time in column (7).

The effect of the number of unique study tracks declines when included jointly with travel time to the nearest school as shown in columns (5)–(9). In the models without municipal and school fixed effects, the estimate is, however, close to significant at 10% level, and in the fixed effects models neither of the two measures of geographical constraints is significant.[20] These results suggest that proximity to one school with 5–8 study tracks is better than proximity to 2–4 schools with only one study track each. Competition between schools seems to be of less relevance than a large choice set of study tracks close to home.

The models presented in Table 4 use a variable that counts the number of unique study tracks within 30 min travel time. Using 30 min as a travel time threshold is somewhat ad hoc. Using Fig. 5, we investigate whether the results are robust to other thresholds.[21] We initially use 3 min as the travel time threshold, and the subsequent estimates add 3 min to the previous specification.[22] The solid line in Fig. 5 shows estimated coefficients for the number of unique study tracks variables corresponding to the different threshold definitions, while the dotted lines show ±2 standard errors. As expected, the estimated impact is small for travel time thresholds close to zero. The estimated effect increases up to the 30 min threshold and is relatively stable thereafter. Choices that require travel time exceeding 30 min do not seem to improve the probability of on-time graduation.

5.4. Heterogeneous effects

An interesting question is to what extent the impact of geographical constraints differs between student groups. Dickerson and McIntosh (2010) find that the effect of distance on upper secondary school choices in the U.K. is heterogeneous with significant effects only for students with mediocre prior achievement and students with disadvantaged family background. Table 5 reports results for travel time to nearest schools for different groups according to GPA from lower secondary education, gender, parents' education, and immigration status. The table presents results for the model specifications both with and without fixed effects.

The three first columns in Table 5 show the effects in different parts of the ability distribution measured by GPA quartile. While the estimated effect of travel time is statistically insignificant and close to zero in the fourth quartile, it is clearly significant and numerically quite strong in the second and third quartile. For the first quartile,

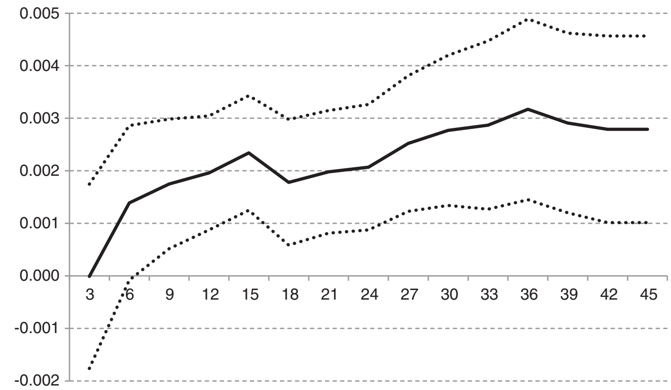

Fig. 5. The effect of proximity to unique study tracks, different thresholds of travel time in minutes. Estimates ±2 standard errors.

the effect is numerically similar to the average effect, but clearly insignificant in the fixed effects specification. This pattern seems sensible, since students with high grades (low grades) have very high (low) probability of graduating on time anyway. Thus, geographical constraints seem to be important only for students at the margin of graduating. For a student in the second or third quartile, the estimated coefficients imply that increasing the travel time to the nearest school by 30 min decreases the probability of graduating on time by about 3.5 percentage points.

The next two columns split the sample by gender. The estimated coefficients indicate that the impact of geographical constraints is similar for boys and girls. Columns (6)–(8) split the sample according to parents' education. The effect of travel time seems to be largely independent of parental education. Columns (9) and (10) split the sample with respect to immigration status. The effect of geographical constraints are insignificant for immigrants, but with relatively large point estimates. The estimated effects on non-immigrants are comparable to the average effect for the whole sample. The imprecise estimates for the immigrant sample can be a small sample property. In addition, since most of the immigrants are located in the main cities, the variation in our measures of geographical constraints is low in the immigrant sample.

5.5. Differences-in-differences approach

A causal interpretation of the estimated negative effect of travel time in the fixed effects model rests on the assumption that residential sorting within the municipalities is random. Foley et al. (2009) finds that parents' self-reported valuation of education is a strong predictor for high school dropout in Canada. If parents with low unobserved valuation of education tend to settle in the periphery of the municipalities, our previous results will be biased. By comparing results by sub-groups, such sorting can be differenced out of the model.

Column (1) in Table 6 replicates columns (1)–(3) in Table 5 in an interaction term framework with students in the fourth GPA quartile as a reference category. Thus, the level effect of travel time estimated is the effect for students in the fourth quartile of the distribution of GPA from lower secondary education. The interaction terms test whether the effect of travel time in the other quartiles differs from the fourth quartile, in contrast to the tests presented in Table 5, where the null is zero impact.

The rest of Table 6 includes ward fixed effects and represents the differences-in-differences approach. In this approach, the level effect of travel time to the nearest school cannot be identified, but we can test whether the effect differs for students in different quartiles of the GPA distribution. Thus, the numerical impact of the interaction terms is directly comparable to the estimates in column (1). The

[20] Notice, however, that travel time to nearest school and the number of unique study tracks within 30 min travel time are jointly significant in columns (8) and (9) in Table 4 (p-values of 0.04 and 0.08, respectively).

[21] When calculating the number of unique study tracks for different travel time thresholds, the number of matches between ward midpoints and schools increases when the threshold level increases. The 3 min threshold yields 5600 matches and the 45 min threshold yields 225,000 matches.

[22] The model used in these functional form robustness checks is in other respects equal to that in column (3) in Table 2.

Table 5
Heterogeneous effects on on-time graduation.

	GPA in 1 quartile	GPA in 2 or 3 quartile	GPA in 4 quartile	Girls	Boys	None of the parents have upper sec. education	At least one parent has upper sec. education	At least one parent has at least bachelor degree	Immigrant	Non-immigrant
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A										
Travel time to nearest upper secondary school, hours	−0.040*	−0.068*	−0.006	−0.049*	−0.042*	−0.063*	−0.046*	−0.038*	−0.088	−0.045*
	(−2.19)	(−3.63)	(−0.27)	(−3.10)	(−2.21)	(−3.45)	(−2.68)	(−2.00)	(−1.25)	(−3.45)
Municipality FE	No	No	No	No	No	No	No	No	No	No
Upper Secondary school FE	No	No	No	No	No	No	No	No	No	No
Panel B										
Travel time to nearest upper secondary school, hours	−0.038	−0.073*	0.023	−0.020	−0.047	−0.094	−0.037	−0.038	−0.160	−0.034
	(−0.81)	(−2.47)	(0.93)	(−0.79)	(−1.52)	(−1.68)	(−1.35)	(−1.14)	(−0.78)	(−1.83)
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Upper Secondary school FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	13,000	24,848	13,636	24,952	26,532	7,567	24,220	19,697	2,758	48,726
Mean on-time graduation	0.173	0.584	0.880	0.636	0.486	0.348	0.509	0.701	0.478	0.556

Note: Same model specification as in column (3) in Table 2 except as indicated. t-values based on standard errors clustered at the regional level in parentheses.
* Denotes statistical significance at 5% level.

heterogeneous effects estimated are in fact larger than in the model without ward fixed effects. Travel time has a significantly larger effect for students with mediocre prior achievement than for students in the fourth quartile, while the difference is not significant for students in the first quartile. The model in column (3) drops the latter interaction term, without affecting the importance of travel time for mediocre students.

Columns (4) and (5) in Table 6 account for upper secondary school quality and other unmeasured school level effects by including fixed effects for the school in which the student enrolls after graduation from lower secondary education. The school and municipality fixed effects are interacted with the GPA quartile. The variation exploited for identification is across students attending the same upper secondary school living in the same ward, but belonging to different quartiles of the GPA distribution. Qualitatively, the model results do not change; travel time is most important for students with mediocre prior achievement.

An indication of whether this strategy successfully accounts for the unobservable characteristics relevant for sorting is whether the result is sensitive to the inclusion or exclusion of observable characteristics. Column (6) in Table 6 excludes all control variables from the model, except the indicators for GPA quartile. Then the estimated effect in fact decreases to −0.060.[23] The fact that the numerical effect is largest when control variables are included suggests that it is not biased towards zero by omitted variables. Taken together, the estimates in Table 6 imply that increasing travel time by 30 min reduces the probability of graduating by 3–6 percentage points more for students with mediocre prior achievement than for other students.

[23] This result indicates that the variable of main interest, the interaction between mediocre prior achievement and travel time, is only weakly correlated with the control variables. We have performed some "balancing tests" to investigate the correlations between the interaction variable of interest and the individual characteristics. The conditional correlations are calculated by regressing the interaction variable of interest on ward fixed effects, an indicator for the relevant GPA quartile, and the individual characteristic of interest. Out of the 16 conditional correlation coefficients, none is significant at 1% level, two are significant at 5% level (parent with master degree and benefits due to disease before age 18), and two are significant at 10% level (first generation immigrant and divorced parents). The other 12 correlations are insignificant.

It is still possible to imagine that the estimates are biased due to sorting in the housing market. Unobservable factors such as perceived valuation of education might be more important determinants of residential location for parents of students with mediocre achievement than for other parents. Although it may seem unlikely that such mobility drives the results, because relatively few parents move while their children are in compulsory education, we can explore this issue a bit further. We have information on whether the students have moved across municipalities during compulsory education. Table 1 shows that 87% of the students did not move. Excluding the "mobile" students and parents from the sample, the estimates are of the same magnitude as those presented in Table 6. For the model specification in column (3) and (6), the estimated effects are −0.070 and −0.077, respectively. Overall, the results for the differences-in-differences approach provide evidence that the fixed effects results above are not severely biased due to within-municipality sorting.

6. Concluding remarks

Motivated by conventional human capital theory and the frequent use of distance to colleges and schools as instruments for educational attainment in Mincerian earnings equations, this paper offers a detailed analysis of the relationship between graduation from upper secondary education and geographical constraints facing students. Using Norwegian data, we find that increased real travel time from parent's home to the nearest upper secondary school decreases the probability of graduating on time. We also find that a higher number of unique study tracks within 30 min commuting distance from the parent's home have a significant positive impact on the propensity to graduate on time. These results hold for a number of different identification strategies, including detailed fixed effects model, IV-models, and differences-in-differences models. Our results thus support the use of geographical constraints as instruments for educational attainment in earnings equations. However, they also suggest that the results from such IV-models must be interpreted with care, because we find that the effects of geographical constraints are not overwhelmingly

Table 6
Differences-in-differences estimates. Dependent variable is on-time graduation.

	(1)	(2)	(3)	(4)	(5)	(6)
Travel time to nearest upper secondary school, hours	−0.006 (−0.27)	–	–	–	–	–
Interaction between travel time and						
GPA in 1 quartile	−0.034 (−1.27)	−0.063 (−1.75)	–	−0.082 (−1.19)	–	–
GPA in 2 or 3 quartile	−0.063* (−2.53)	−0.096* (−3.10)	−0.065* (−3.08)	−0.129* (−2.38)	−0.090* (2.02)	−0.060* (−5.60)
Ward FE	No	Yes	Yes	Yes	Yes	Yes
Municipality FE	No	No	No	Yes	Yes	No
Upper secondary school FE	No	No	No	Yes	Yes	No
Control variables interacted with GPA quartile	Yes	Yes	Yes	Yes	Yes	No

Note: Column (1) replicates the results in columns (1)–(3) in Panel A in Table 5. All variables, except ward fixed effects but including the other fixed effects, are interacted with GPA quartiles. Ward fixed effects are assigned according to residence in January 1, 2002. Municipality fixed effects interacted with the GPA quartiles are included in columns (4) and (5) in order to make the models equal to the models in Panel B in Table 5 in all aspects except the ward fixed effects. 51,484 observations. t-values based on standard errors clustered at the regional level in parentheses.
* Denotes statistical significance at 5% level.

strong, and they are clearly heterogeneous. In particular, we find that the impact is most prevalent among students with mediocre prior achievement. Geographical constraints seem to mostly affect the graduation probability of students at the margin of graduating.

Combining our results with recent evidence on the importance of human capital for regional economic growth, geographical constraints on schooling opportunities may have lasting impact on regional income differences. One may be tempted to take the negative impact of geographical constraints on educational attainment found in this paper as support for policies that decentralize educational institutions. However, such an interpretation is premature since structural policy recommendations must take economies of scale in education production into account. Decentralized systems might be costly compared to other policies to foster human capital investments and regional economic growth.

Acknowledgement

Excellent research assistance from Ole Henning Nyhus and comments from Jørn Rattsø, Arnt Ove Hopland, participants at the annual congress of the European Economic Association in Oslo 2011, the annual conference of the European Association of Labour Economists in Cyprus 2011, the annual meeting of Association of American Geographers in New York 2012, a seminar in Trondheim, the two reviewers, and the editor are greatly acknowledged.

Appendix 1. Definitions of explanatory variables and data reduction

Individual and family variables

GPA: Average grade from lower secondary school

Girl: Binary explanatory variable equal to 1 if student is a girl and 0 if student is a boy.

Both parents have compulsory education only: Binary variable, registered in the year the student was 16 years old.

At least one parent has some upper secondary education: Binary variable, registered in the year the student was 16 years old.

Table A1
Data reduction.

	Observations	Percent of population
Population graduated from lower secondary education in 2002	55,793	100
Not born in 1986	2889	5.2
Missing ward identifier	329	0.6
Missing information on GPA	1091	2.0
Regression sample	51,484	92.3

At least one parent has a bachelor degree: Binary variable, registered in the year the student was 16 years old.

At least one parent has a master's or Ph.D. degree: Binary variable, registered in the year the student was 16 years old.

First generation immigrant: Binary variable equals 1 if student born abroad with both parents born outside Norway, 0 otherwise.

Second generation immigrant: Binary variable equals 1 if student born in Norway, with both parents born outside Norway, 0 otherwise.

Explanatory variables measured at the ward level

GPA average

Share of students with both parents having only compulsory education

Share of students with at least one parent with some upper secondary education

Explanatory variables measured at the municipal level

Square km per inhabitant in municipality

Share of population living in municipality's rural areas

Appendix 2. Instrumental variable estimates. Complete results

Table A2
Complete results for models in columns (4)–(6) in Table 3.

	(1)	(2)	(3)
	Reduced form on-time graduation	Reduced form for travel time (1. stage)	Structural model for on-time graduation
Travel time to nearest upper secondary school, hours	–	–	−0.061* (−2.56)
Share of population living in rural areas in municipality	−0.0095 (−0.75)	0.32* (6.54)	–
Square km per inhabitant in municipality	−0.033* (−2.28)	0.35* (3.36)	–
GPA	0.28* (83.0)	0.0017 (1.91)	0.28* (83.4)
Girl	0.027* (4.28)	−0.0023 (−1.65)	0.026* (4.32)
First generation immigrant	0.023 (1.53)	−0.0026 (−1.02)	0.023 (1.54)
Second generation immigrant	0.038* (3.92)	0.0043 (0.98)	0.038* (3.99)
Birth month	0.0039* (7.71)	0.00036 (1.90)	0.0039* (7.84)
Parents married	0.058* (12.4)	0.0064* (3.11)	0.059* (12.6)
Parents divorced	0.013* (2.22)	−0.0041 (−1.42)	0.013* (2.21)
At least one parent upper secondary education	0.020* (3.19)	0.0041* (3.40)	0.021* (3.28)
At least one parent bachelor degree	0.031* (4.80)	−0.0034* (−2.16)	0.031* (4.82)
At least one parent master degree	0.028* (2.71)	−0.014* (−3.87)	0.027* (2.65)

(continued on next page)

Table A2 (*continued*)

	(1) Reduced form on-time graduation	(2) Reduced form for travel time (1. stage)	(3) Structural model for on-time graduation
Parental income in quartile 2	0.013*	−0.00067	0.012*
	(2.11)	(−0.29)	(2.13)
Parental income in quartile 3	0.025*	−0.0034	0.025*
	(4.40)	(−1.26)	(4.42)
Parental income in quartile 4	0.028*	−0.0048	0.028*
	(4.65)	(−1.76)	(4.64)
Student moved between municipalities in age 6 16	−0.034*	−0.0021	−0.034*
	(−5.07)	(−0.76)	(−5.15)
Student mobility unknown	−0.0047	−0.0047	−0.0049
	(−0.33)	(−0.75)	(−0.35)
Benefits due to disabilities before age 18	0.0083	0.0094	0.0087
	(0.54)	(1.12)	(0.57)
Benefits due to disease before age 18	−0.059*	0.0036	−0.059*
	(−3.86)	(0.75)	(−3.89)
Ward has rural settlement	−0.0063	0.094*	0.00030
	(−1.13)	(12.1)	(0.048)
Type of settlement in ward unknown	0.024	0.13*	0.033
	(0.74)	(5.54)	(1.00)
Log (Number of students in same cohort in the same ward)	0.0022	−0.0086*	0.0017
	(0.76)	(−2.15)	(0.57)
Average GPA in the ward	−0.011*	0.0064	−0.010*
	(−2.55)	(1.19)	(−2.52)
Average share of parents with at least upper secondary education in the ward	0.023	−0.052*	0.020
	(1.88)	(−3.21)	(1.64)
Average share of immigrants in the ward	−0.012	−0.066*	−0.016
	(−0.60)	(−2.75)	(−0.78)
Region fixed effects	Yes	Yes	Yes
Upper secondary school fixed effects	Yes	Yes	Yes
Observations	51,484	51,484	51,484
R-squared	0.349	0.887	0.349
F-test for weak instruments	–	60.7	–
Test for overidentifying restrictions (p-value)	–	–	0.265

Note: t-values based on standard errors clustered at the regional level in parentheses.
* Denotes statistical significance at 5% level.

References

Andersson, R., Quigley, J.M., Wilhelmsson, M., 2004. University decentralization as regional policy: the Swedish experiment. Journal of Economic Geography 4, 422–448.

Andersson, R., Quigley, J.M., Wilhelmsson, M., 2009. Urbanization, productivity, and innovation: evidence from investment. Journal of Urban Economics 66, 2–15.

Becker, G., 1964. Human Capital: a Theoretical Analysis of Special Reference to Education. Columbia University Press, New York.

Becker, S.O., Hornung, E., Woessmann, L., 2011. Education and catch-up in the industrial revolution. American Economic Journal: Macroeconomics 3, 92–126.

Becker, S.O., Siebern-Thomas, F., 2007. Schooling infrastructure, educational attainment and earnings. Mimeo.

Becker, S.O., Woessmann, L., 2009. Was Weber wrong? A human capital theory of Protestant economic history. Quarterly Journal of Economics 124, 531–596.

Belley, P., Lochner, L., 2007. The changing role of family income and ability in determining educational achievement. Journal of Human Capital 1, 37–89.

Ben-Porath, Y., 1967. The production of human capital and the lifecycle of earnings. Journal of Political Economy 75, 352–356.

Black, D.A., McKinnish, T.G., Sanders, S.G., 2005. Tight labor markets and the demand for education: evidence from the coal boom and bust. Industrial & Labor Relations Review 59, 3–16.

Bradley, S., Lenton, P., 2007. Dropping out of post-compulsory education in the UK: an analysis of determinants and outcomes. Journal of Population Economics 20, 299–328.

Cameron, S.V., Taber, C., 2004. Estimation of educational borrowing constraints using returns to schooling. Journal of Political Economy 112, 132–182.

Card, D., 1995. Using geographic variation in college proximity to estimate the return to schooling. In: Christofides, L.N., Grant, K.E., Swidinsky, R. (Eds.), Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp. University of Toronto Press, pp. 201–222.

Carneiro, P., Heckman, J.J., Vytlacil, E., 2011. Estimating marginal returns to education. American Economic Review 101, 2754–2781.

Ciccone, A., Papaioannou, E., 2009. Human capital, the structure of production, and growth. The Review of Economics and Statistics 91, 66–82.

Clark, D., 2011. Do recessions keep students in school? The impact of youth unemployment on enrolment in post-compulsory education in England. Economica 78, 523–545.

de la Fuente, A., 2002. On the sources of convergence: a close look at the Spanish regions. European Economic Review 46, 569–599.

de la Fuente, A., Vives, X., 1995. Infrastructure and education as instruments of regional policy: evidence from Spain. Economic Policy 10, 11–51.

Dickerson, A., McIntosh, S., 2010. The Impact of Distance to Nearest Education Institution on the Post-Compulsory Education Participation Decision, Sheffield Economic Research Paper Series, p. 2010007.

Do, C., 2004. The effects of local colleges on the quality of college attended. Economics of Education Review 23, 249–257.

Eckstein, Z., Wolpin, K.I., 1999. Why youths drop out of high school: the impact of preferences, opportunities and abilities. Econometrica 67, 1295–1339.

Falch, T., Strøm, B., 2011. Schools, ability and the socioeconomic gradient in education choices. CESifo Working Paper No.3313.

Foley, K., Gallipoli, G., Green, D.A., 2009. Ability, parental valuation of education and the high school dropout decision. IFS Working Papers W09/21.

Frenette, M., 2009. Do universities benefit local youth? Evidence from the creation of new universities. Economics of Education Review 28, 318–328.

Gibbons, S., Vignoles, A., 2011. Geography, choice and participation in higher education in England. Regional Science and Urban Economics 42, 93–113.

Griffith, A.L., Rothstein, D.S., 2009. Can't get there from here: the decision to apply to a selective college. Economics of Education Review 28, 620–628.

Heckman, J.J., LaFontaine, P.A., 2010. The American high school graduation rate: trends and levels. The Review of Economics and Statistics 92, 244–262.

Kjellström, C., Regnér, H., 1999. The effects of geographical distance on the decision to enrol in university education. Scandinavian Journal of Educational Research 43, 335–348.

Kling, J.R., 2001. Interpreting instrumental variables estimates of the returns to schooling. Journal of Business and Economics Statistics 19, 358–364.

Koedel, C., forthcoming. Higher education structure and education outcomes: Evidence from the United States. Education Economics doi:10.1080/09645292.2011.616714.

Maani, S.A., Kalb, G., 2007. Academic performance, childhood economic resources, and the choice to leave school at age 16. Economics of Education Review 26, 361–374.

OECD, 2011. Education at a glance. http://www.oecd.org/dataoecd/61/2/48631582.pdf2011.

Rice, P., 1999. The impact of local labour markets on investment in further education: evidence from the England and Wales youth cohort studies. Journal of Population Economics 12, 287–312.

Sa, C., Florax, R.J.G.M., Rietveld, P., 2006. Does accessibility to higher education matter? Choice behaviour of high school graduates in the Netherlands. Spatial Economic Analysis 1, 155–174.

Van Ommeren, J.N., Gutiérres-i-Puigarnau, E., 2011. Are workers with a long commute less productive? An empirical analysis of absenteeism. Regional Science and Urban Economics 41 (1), 1–8.

Zenou, Y., 2002. How do firms redline workers? Journal of Urban Economics 52, 391–408.