

Supply of Education and Schooling choices

Hanna Virtanen*, Lotta Väänänen**

** Aalto University, School of Economics, HECER, FDPE ** University of Mannheim*

Hanna Virtanen, hanna.virtanen@hse.fi, Aalto University, School of Economics, P.O.Box 21240, FI-00076 AALTO

We examine how the regional supply of education affects the schooling choices of young people. We use rich individual-level data on 4 cohorts of Finns leaving comprehensive school in 2000-2003 to study the sorting of individuals across 8 schooling alternatives (high school and 7 vocational fields). We use discrete choice models with random coefficients, allowing for heterogeneous preferences. We find that the distance to the nearest educational institute offering an alternative has a statistically significant, although a small effect on which schooling alternative is chosen. The estimated affects are larger for boys. We also find, similar to previous studies that family background (namely parents' education and socioeconomic status) and previous school performance play a role in influencing the schooling choice.

Keywords: upper secondary education; vocational education, schooling choices; supply of education; random coefficients

We would like to thank Pekka Ilmakunnas and Otto Toivanen for comments. We also thank Satu Nurmi at Statistics Finland for her work in constructing the data. We would like to thank the Yrjö Jahnsson Foundation for financial support. Väänänen also gratefully acknowledges financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15. The usual caveat applies.

1 Introduction

A vast literature has established that education, by means of increasing human capital, has a significant influence on the productivity and success of individuals, firms, and societies. Hence, understanding the schooling decisions of individuals is important. In this paper, we examine how the regional supply of education affects the schooling choices of young people, using rich data on the application and enrolment decisions of four complete cohorts of Finns leaving compulsory schooling in the years 2000-2003. We study the sorting of individuals across schooling alternatives, namely high school and 7 fields of vocational education. The choice of the upper secondary schooling alternative is a crucial step in defining the career opportunities available in later stages. Our data allows us to thoroughly analyze the individuals' schooling decisions, taking into account all the schooling alternatives available and their features, while controlling for detailed family background information, school performance, characteristics of the comprehensive schools, and regional characteristics. Supply side characteristics that measure the regional availability of the various alternatives in upper secondary education are of particular interest, as these are factors that are determined by educational policy.

Supply side factors have been shown to play a role in determining the schooling decisions of individuals. One supply factor is the location of educational establishments, which generates variation in individuals' proximity to schooling. Distance to schooling, in turn, can affect the costs of schooling in several ways: direct financial costs of re-allocation or commuting, emotional costs associated with leaving home, and information costs when seeking information on the schooling options. Besides the distances to the schooling alternatives, the availability of open schooling positions may also affect the schooling decision, for instance via increased competition. In its recent definition of policy, the Finnish Committee for Education and Culture stated that the target of the education policy in Finland for the 2012 is to diminish the regional differences in the supply of education (the number of open positions per cohort).¹ If the supply of education has an impact on the schooling decisions, then possible regional disparities in the availability of education can introduce inequalities in human capital accumulation between individuals living in different regions. A poor availability of education may also be linked to social exclusion, as participation in upper-secondary schooling is considered to be crucial to the later success of the individuals. Availability of schooling may also have larger effects on some sub-groups of individuals (low income families etc.) and thus, a deficiency in the local supply of education may further increase the inequality of these groups. Regional differences in the accessibility of education may also boost differences in skill composition between regions.

¹ "Sivistyspoliittinen ministeriryhmä linjasi koulutuksen alueelliset tavoitteet vuodelle 2012." 16.5.2008.

There are some studies that examine the effects of supply on the schooling decision (e.g. Gibbons & Vignoles, 2009; Frenette, 2004; 2006; Spiess & Wrohlich, 2008).² These studies focus on the effects of supply on the participation decision and on higher education. In addition to its effect on participation, supply may have an effect on the schooling alternative chosen. We contribute to the literature by examining the effects of supply on the sorting of individuals across schooling alternatives. In addition, to our knowledge, we are first to examine the effects of the availability of schooling on the upper-secondary schooling decisions, instead of the schooling decision of higher education examined in the previous papers. The availability of schooling may be of particular importance to the 16-year old individuals graduating from compulsory school as they are less likely to migrate to another region than the older age groups. Finally, our data also allows us to control for most of the factors found important in the previous literature.

We study how the availability of schooling affects the schooling decisions of students leaving compulsory school, namely the choice among fields of studies in vocational education and high school. We use a rich cross section data on four complete cohorts of individuals graduating from the 9th grade in 2000-2003. We use a conditional logit model to model the discrete choice among the alternatives, and use a random coefficients specification to allow for heterogeneous effects of the supply variables. The variables we use to describe the supply of schooling are (i) the distance to the nearest municipality offering each alternative, (ii) an indicator for whether an alternative is offered in municipality, and (iii) the number of open positions in each alternative per size of the cohort in a region.

Our results show that the availability and proximity of a schooling alternative has a positive and statistically significant role in influencing the decision to choose that alternative. Furthermore, the cross-effects of the proximity of other fields are negative. For boys, the estimated effect is larger than for girls, and we also find significant heterogeneity of the effect (i.e. significant standard deviation of the random coefficient).

Our study also provides interesting insights into the effects of family background and the individuals' previous school performance on the schooling choices. We find that the parents' fields of education play a significant role in influencing the schooling decision. This indicates that there is intergenerational persistence in the fields of education. In particular the father's educational choices are influential, and especially for boys. In addition, also the parents' level of education, and, to some extent, their socioeconomic status has the expected effect. Finally, the previous school performance of the individual has an important effect on the schooling choice.

The rest of the paper is organized as follows. We begin with a brief literature review. In Section 3 we describe the schooling alternatives and the application system in upper secondary

² Supply characteristics are often used as instruments for the schooling decision (to study the returns to education or other educational outcomes), as they are perceived to be uncorrelated with the studied outcomes of the education (see Card 2001 for a survey).

education in Finland. In section 4 we describe the data in detail. In section 5 we present the empirical analysis and the results, and in section 6 the conclusions.

2 Literature review

Most of the previous literature on educational choices focuses on the level of educational attainment (i.e. the participation decision in higher education, the number of years of schooling etc.). Literature on the participation decision in higher education shows that factors such as previous school performance, parental background, peers, and the economic situation of the individual and family, are important for individuals' schooling decisions. The effects of the distance on the education attainment have been explored in more detailed only in few studies.

Spieß & Wrohlich (2008) study how the distance to nearest public university affects the enrolment decision of individuals who have passed their university entrance exams (Abitur). They use a subsample of 1,223 individuals from the German Socio-Economic Panel for the years 1993 to 2006 for whom they are able to control for several family characteristics (such as parents' level of education and annual net income) and regional characteristics (urbanization and population density). Based on a discrete time hazard rate model, they show that one kilometre distance decreases the probability to participate in higher education by 0.2-0.3 percentage points. They do not find any differences in the results for student from families with lower income or lower level of education. Frenette (2004) uses a multinomial logit model to examine whether the geographical accessibility of universities and colleges have an effect on the individuals decision to participate in a local college, university or not to study in higher education. They include dummies for year and province and controls for the parent's education level and income. They conclude that having a university within commuting distance (less than 80 kilometres) increases the probability of attending university significantly and decreases the probability of attending a local college. The estimated effects are of a larger magnitude for lower income families. These studies do not include any controls for previous school performance or other individual characteristics (besides gender) which might effect the enrolment decision.

Gibbons & Vignoles (2009) study how the distance to university affects the participation decision and choice of university in the UK by estimating logit models of participation and conditional/multinomial models of institution of choice. They control for detailed information on the individuals and regions while including a full set of institutional dummy variables. They do not, however, have any information on parental education. They show that the distance has only a small affect on the decision to participate in higher education but a large effect on the choice of a university. The effect of distance on the institute of higher education chosen is particularly strong for Bangladeshi and Pakistani girls and low income students. (The results by Gibbons & Vignoles (2009) and Frenette (2004) indicate that there might be some substitution effects with the distance.)

The schooling decision is a more complex phenomenon than choosing the optimal level and recently there has been some research on the choice between fields of education, focusing on the effects of factors such as parents' education (Dustmann 2004) and expected earnings (Beffy et. al. 2009, Arcidiacono 2004). Beffy et al. (2009) examine the effect of expected earnings on the choice of major and find that expected earnings have a quantitatively small, yet significant, impact on the choice of major. Rather, the choice is related both to schooling preferences and abilities, indicating that nonpecuniary factors are a key determinant of schooling choices. Similarly Arcidiacono (2004), in a dynamic model of college and major choice, finds that most of the ability sorting to majors is due to preferences rather than monetary rewards. Dustmann (2004) focuses on the intergenerational transmission of education by examining parental factors in influencing the schooling decision. He shows that educational mobility is limited as parental factors play a large role in determining the choice of secondary school track, which further influences the post-school educational choices and ultimately the income of the individual. These papers have not included supply-side features of the educational system in the analysis. Our paper therefore complements their analysis by including in the same analysis supply-side characteristics as well as detailed information on individuals and parents, and allowing for unobserved idiosyncratic preferences and random coefficients.

Most of the previous literature has examined the post-secondary schooling decision (participation in higher education, choosing a major etc.). There are a few exceptions (Dustmann 2004, Beffy et. al. 2009). Studying the schooling decisions of individuals at the time of leaving comprehensive schooling is particularly interesting, as these decisions involve a choice of the direction of career the individuals start investing in and the kind of human capital they begin to acquire. In Finland, people complete their compulsory schooling at the age of 16 and then choose between high school, which usually leads to higher education, and vocational education, where they also choose among different fields of study.

3 The upper secondary education system in Finland

In the following sections, we describe the upper secondary schooling system in Finland. We start by describing the available schooling alternatives and the application and student selection processes. We also show some descriptive statistics on the choices of the young people in our data. In Section 3.2 we overview the administration and implementation of education policy, which underlies the supply of upper-secondary education. We also present evidence on the regional variation in the supply of education based on our data.

3.1 *Schooling alternatives*

Compulsory education in Finland consists of nine years of comprehensive school and it typically ends at the age of 16. There is also a possibility to continue for one additional year (10th grade) in comprehensive schooling after finishing the 9th grade. After completing

comprehensive schooling, one can continue to upper secondary education or decide not to study further. It is strongly recommended that young people continue studying after compulsory schooling and one of the objectives of education policy is to provide upper secondary education to all of each age group free of charge.

Upper secondary education is divided into general and vocational education. General education (from now on high school) continues the educational tasks of the comprehensive school. The scope of the syllabus in high school is three years³ and it ends with matriculation. Vocational education and training includes 7 fields of education, which contain 30 subfields, which again include about 50 different vocational qualifications and about 120 studying programs.⁴ Vocational upper secondary qualification can be obtained through a school-based education system⁵, through apprenticeship training⁶, or through a competence test⁷. The majority of those who enter vocational education directly after compulsory school attend the school-based system. The scope of the upper secondary vocational qualification taken after comprehensive school is three years.⁸ Some of the studies in both high school and vocational education are accredited, and it is also possible to attain the matriculation and a vocational qualification simultaneously. This usually means that the studies take 4 years⁹.

Both the matriculation examination and the vocational qualification with a scope of at least three years provide general eligibility for universities and polytechnics. However, the majority of new students in higher education have completed the matriculation examination. According to Statistics Finland (Kaukonen, 2009), during the first few years after graduation, less than 15 per cent of those who graduate from vocational education continue to higher education. The corresponding figure for those who graduate from high school is more than 70 per cent. These shares have stayed consistently the same over the past decade.

From our data, we see that more than 90 per cent of the cohorts in the 9th grade continue to upper secondary education each year (2000-2003), and about 2 per cent to the additional 10th grade. Over 5 per cent do not immediately continue in education. About 60 per cent of those

³ General education can be accomplished in two, three or four years.

⁴ See appendix A for more detailed information about the subfields in vocational education.

⁵ The education takes mostly place in a vocational institution, but this form of vocational education also includes approximately six months of on-the-job-learning.

⁶ Most of the education takes place at the workplace, but there are also some courses arranged in the institutions.

⁷ Competence tests are independent of how the vocational skills have been acquired.

⁸ There are also a few vocational qualifications, which have a scope of less than three years. These options are rarely chosen directly after compulsory education and they do not provide eligibility for further studies.

⁹ In the official statistics students are usually assigned to one education institution at a time and thus, we do not observe this in our data.

who continue to upper secondary education go to high school and the rest to vocational education. More than 99 per cent of those starting in vocational education are enrolled in the school-based education system. These figures are also consistent with those published annually by Statistics Finland. According to Statistics Finland, the share of vocational education has been increasing and the share of young people who remain outside of the upper secondary education has been decreasing over the last decade.

Admission to the upper secondary education takes place through the joint application system maintained by the Finnish National Board of Education. Most of the positions in the secondary education are filled through this system¹⁰. Individuals can simultaneously apply to five different schooling positions. A schooling position is defined as an education institution-schooling alternative combination. Usually the schooling alternative is defined at the level of the vocational qualification (or high school), but it can also be more specific (studying program) or more general (sub-field). Thus, students can apply to the same schooling alternative in five different education institutions or to five different schooling alternatives in the same education institution, for instance. Individuals rank their requests and receive one offer at a time. If a person does not gain admittance to the school of his or her first choice, the other requests are considered in their order of ranking. The Ministry of Education determines the general student selection criteria, but an education provider can decide on the specifics of the student admission criteria (e.g. use of entrance or aptitude tests). For those positions that do not get filled at the first application, a replacement application process is organized.

3.2 *Supply of upper secondary education*

Upper secondary education is provided by local authorities, municipal consortia or other organisations authorised by the Ministry of Education. According to the statistics of educational institutions by Statistics Finland, there were about 440 high schools and around 200 vocational education organisations in the year 2004¹¹. The licence to provide upper secondary education (*järjestämislupa*) defines the maximum number of students an education organisation is allowed to have per field of education.

One of the objectives of the educational policy is to guarantee educational basic rights to all young people regardless of where they live. However, Finland is a sparsely inhabited and geographically large country, and our data shows that substantial differences exist in the availability of the various schooling alternatives between the regions. Table 1 shows the

¹⁰ There are some types and fields of education, which do not use the joint application system (e.g. the smaller scale vocational qualifications, vocational qualifications in specialized fields such as music and dance)

¹¹ The number of education organizations is probable a little higher during our observation period as the amalgamation of education organizations has been ongoing trend in upper secondary education during the last decade.

municipality-level descriptive statistics on the availability of each alternative during our observation period 2000-2003. Columns 1-4 show the mean, maximum and standard deviations of distances to the nearest municipality offering the alternatives. The fifth column shows the fraction of municipalities offering each alternative. We see that high school is most prevalent, present in about 70% of the municipalities and the average distance to the nearest municipality offering high school is 8km. However, some municipalities face longer distances to high school, up to a maximum of 74km. Vocational education is more scarcely available and there are also large differences between fields. More than 40% of the municipalities offer at least one vocational education. Of vocational fields, the Technology and Transport sector, and Hotel, Catering and Home Economics sector are most widespread, offered in about 30% of the municipalities, whereas the Humanist and Teaching sector is present only in about 7% of municipalities.

Columns 6-8 in Table 1 show the degree of within-municipality variation in the availability of the alternatives. There is no within variation in the presence of high schools, and also only little variation in the presence of vocational fields. We can also see a tendency for a reduction in supply, with more counties ceasing to provide some educational fields than establishing new.

[Table 1 here]

4 Data and descriptive analysis

We have a very detailed individual level dataset on four full cohorts of Finns leaving comprehensive school in the years 2000-2003. We observe the application and enrolment decisions into vocational schools and high schools, important personal and family background characteristics, as well as measures of supply of education in each municipality. We start by describing our data in Section 4.1. We describe the various data sources used to assemble the dataset and the construction of the sample. We then discuss the variables we use to measure individuals' schooling choices, the regional supply of education, the characteristics of the schooling alternatives, and the various individual and parents' characteristics. In Section 4.2 we present a descriptive analysis of the schooling choices and their distributions conditional on the supply as well as various other characteristics.

4.1 Data

4.1.1 Data sources and sample

We use a rich dataset on four cohorts of individuals leaving the compulsory school during the years 2000-2003. The sample is based on the Application Register of the Finnish National Board of Education and includes those individuals who are in the 9th grade in the spring of one

of the examined years, obtain a leaving certificate during that year¹², and apply to upper secondary education in the joint application that year. We exclude from our sample those individuals who are enrolled in the 9th grade or in the 10th grade in the following year.

We use data sources of Statistics Finland to get detailed information on the schooling decisions of the individuals in the sample. We observe the individuals' enrolment in upper secondary education from the annual data on students and students in apprenticeship training. We exclude from our sample those who are enrolled in special schools or have a special-case student status¹³. Furthermore, we exclude those who are enrolled in apprenticeship training, in vocational education aimed at qualification through a competence test, in specialized school, in folk high schools, or in higher education¹⁴. After these restrictions, we are left with 95 percent of the initial sample and we have about 230,000 individuals in the sample¹⁵.

We complement the dataset with detailed information on the individuals and their parents from the Finnish Longitudinal Employer-Employee data (FLEED) and the Degree Register of Statistics Finland. The same data sources are also used in determining the characteristics of the alternatives and the regional labour market conditions. The supply of education is determined from the online database WERA of the National Board of Education. In addition, we obtain a matrix of inter-municipality driving distances from the Finnish Road Administration.

4.1.2 *Schooling choices*

We determine the outcome variable, namely the schooling choice of the individual, based on the schooling alternative which is assigned to as first request in the joint application. The first request is taken as the individual's optimal schooling alternative given his or her information set (expectations of the success in admission). The enrolment is, instead, determined mutually by the application and selection processes.

From the Application Register we observe first of all, whether the individuals participate in the joint application. For those who apply through the joint application system, we have also information on their application requests (up to five places to which they apply). Data includes information on all the choices (the subfield of education, education organisation and

¹² Individuals who are observed to be in the 9th grade in several years are included in the sample based on the year they receive a leaving certificate. We do not include individuals who are either resident or studying in Åland. We do not observe the inter-municipality distances for these observations.

¹³ The special schools and special student programs are meant for individuals who require additional support in their education (individuals with disabilities etc.).

¹⁴ These education alternatives and their application procedure differ substantially from the other types of schooling alternatives available. Professions in these categories include, for example, fireman and musician.

¹⁵ All the descriptive results in this document are done so far only with a 30 percent random sub-sample (70,691 individuals) of the final sample.

municipality the alternative is located in) and the order of preference which students have assigned to them¹⁶. This provides interesting information on individuals' preferences, e.g. whether they have a strong preference for a particular education (subfield), field or provider.

4.1.3 Supply and alternative specific characteristics

The variables we use to describe the regional supply of education are: i) a dummy variable equal to one if the field of study is offered in the municipality of residence; ii) the distance to the nearest municipality offering the field of study (in km), iii) the number of open positions in the field per cohort in the sub-region (work-in-progress).

We use the classification of municipalities at the year 2007 which contains 400 municipalities each year. We determine the supply of education in these municipalities from the WERA data of the National Board of Education. The data has information on the annual number of open schooling positions announced by the education organisations¹⁷. The open schooling positions are given at the level of education organisation, municipality and education subfield. We aggregate this information to level of education field and municipality. We define the set of open schooling position to include those positions which have the requirement of basic training only in comprehensive schooling, are for young persons and are aimed at completing a degree.

The individual's municipality of residence comes from Application Register and is determined at the time of the application (during the last semester of the 9th grade). We measure the distance from the municipality of residence to the nearest municipality offering each alternative and use this to calculate the minimum distance to each of the alternatives for all the individuals in the sample. If the schooling alternative is offered in the municipality, we define the distance to be 2km¹⁸.

The location of the supply as described above is based on the information of the open positions as announced by the education organisations. (It is likely that) the organisations do not specify the schooling positions they offer for each municipality separately even if they provide education in several municipalities. Hence, we may have some measurement error in the distance to the supply of education. Therefore we use also an alternative way of defining the supply of education (also the only way the supply is defined at the moment). In our data, we have information on the municipality of the education in which students are enrolled. This information is on the basis of the municipality where the schooling takes place and thus, we use it to define an alternative measure of the supply of education in the municipalities. We assume

¹⁶ We aggregate this information to educational field level.

¹⁷ They do not include information on the open positions in apprenticeship training or in education not administrated by the National Board of Education.

¹⁸ The observed minimum of distances between municipalities is 3km.

that if there is supply of a certain education alternative (high school or field of vocational education) in the municipality in the given year, then there is also at least one person from the cohort of that year beginning in this education option during the year. If there are no new students in a schooling alternative in a municipality, then we infer that this alternative is not offered in this particular municipality. This is likely to be quite accurate measure for the municipalities offering each alternative. However, we do not use this measure to examine the effects of the open positions (enrolment of new students) per cohort since the number of new students is determined besides by the availability of open positions, also by the application to these positions, admission of the applicants and acceptance of the offers.

Unlike the previous studies, which have utilised the zip codes of educational institutions (universities) and home, we use the distance between the municipality of supply and residence to define the distance to schooling. Both of these measures introduce a potential source of a measurement error. Even when using zip codes, the distances may be of different effect depending on the means of communication and transport services¹⁹. The distance between municipality of education and home is also imprecise as it gives the same distance to all individuals living in the same municipality and to all education provided within the same municipality. We argue, however, that the distance within a municipality has less of an effect when compared to the distance between municipalities and therefore, our definition of the distance can be considered as an indicative measure of the effective distances. The transportation services are usually better within municipalities as they the municipalities are active in planning, financing and organising public transportation etc. Pupils in comprehensive school are also likely to be better informed about the schooling possibilities in their municipality of residence. Furthermore, the distances between municipalities are also defined based on the driving distances are thus, they provide probably good proxies about the effective distances between municipalities. To diminish the possible measurement error, we use a control for the size of the municipality in square kilometres from the National Land Survey of Finland (not done yet).

To describe the supply environment further, we control for the unemployment rate of youth in the municipality of residence. In addition, we have a categorical variable describing the urbanization of the municipalities, which divide the municipalities into the following three groups: rural, densely populated, and urban. The unemployment rates are calculated from The FLEED data and information on the urbanization of the municipalities comes from regional statistics of the Statistics of Finland. Alternatively/ in addition, we control for the size of the

¹⁹ Frenette (2004) use straight line distances to create an indicator of whether the education institute is “nearby”. Gibbons and Vignoles (2009) define the distances using rail-network distances. Spiess and Wrohlich (2008) use the geographical information system ArcGIS to calculate the distances between the centres of gravity of the zip codes.

municipality in inhabitants from Population Register Centre (not done yet) to capture some differences between the regions which might affect the decision or the effectiveness of the distance.

We have also collected information on the aggregate characteristics of the schooling alternatives from the FLEED data. We use this information to capture the field specific features that might affect the expected discounted reward from each alternative. By using the information on all the Finns between ages 40 to 45 in the year 2000, we calculate the average wages, standard deviation of wages, unemployment rates and proportion of individuals with higher education degree proportion for each schooling alternative. We do the calculations separately for men and women.

4.1.4 Characteristics of the individuals

We have detailed information on the individuals to control for the observed differences between them, which might affect their preferences and the rewards from the alternatives. From the datasets of Statistics Finland, we have personal information on the date of birth, sex, nationality, native language and district of residence, marital status, number and age of children etc. We also have information on the education and working experience of the parents of the individuals in our sample, which is important to control for differences among the individuals arising from family background. Finally, we have information on the comprehensive school the individuals attended and their grades in comprehensive school. We also observe the size of the comprehensive schools as well as the decisions of the school mates.

4.2 Descriptive analysis

We provide descriptive statistics on the supply of education conditional on schooling choices. Schooling choice is measured as the one first indicated in application. We also examine descriptive statistics on the characteristics of the individuals, their parents, and their peers from comprehensive school, conditional on their schooling choices.

4.2.1 Choices and supply of education

Table 2 presents the distribution of the schooling choices in our sample. We can see that there are clear differences in the schooling choices by gender. High school is the most popular choice for both boys and girls, particularly for girls. Of the vocational fields, “Technology and Transport” (35%) is the most common for boys, whereas for girls it is “Social and Health Care Services” (11%).

[Table 2 here]

Table 3 shows the means of minimum distances to each schooling alternative conditional on the individuals’ choices. We can see that those who have chosen “High school”, “Administration and Commerce”, “Culture” or “Humanist and Teaching” have slightly shorter distance to that schooling alternative than what is the proximity on average. However, for other

schooling alternatives this does not hold. It is a matter of regression analysis to determine whether any relationship exists between the proximity of schooling alternatives and individuals' choices.

[Table 3 here]

4.2.2 *Individual choices and characteristics*

In Tables 4 and 5 we provide a comparison between all the schooling alternatives in terms of the characteristics of the individuals who make the choice, characteristics of the comprehensive school they received their diploma from, of the geographic region they live in, and finally a comparison between the characteristics of the alternatives.

In Table 4 we present some evidence on possible peer effects by cross-tabulating the average share of classmates from comprehensive school who choose a given alternative with individuals' choices. We notice a strong tendency for individuals from the same school to make the same choice of alternative in upper-secondary education. These effects of course combine all location and supply effects with potential peer effects.

[Table 4 here]

From Table 5 we can first of all see that the fields of and "Technology and Transport", "Administration and Commerce" and "Natural Resources" are clearly male-dominated, whereas the "Social and Health Care Services", "Humanist and Teaching", "Hotel, Catering and Home Economics" and "Culture" are female-dominated schooling alternatives. We also observe big differences in the average grades of individuals. Expectedly, we see that individuals choosing high school have higher grades in comprehensive school. Of the vocational fields, "Technology and Transport" is the choice for individuals with the lowest average grade in theoretical subjects, indicating that it is an important schooling option of the academically weaker students.

[Table 5 here]

Table 5 provides also the family background characteristics for the individuals conditional on their choice (by first request). We notice a positive correlation between parents' fields of education and the choices of their children. Furthermore, individuals who choose high school have parents who are more educated and have higher incomes. In addition, the parents of the individuals enrolled in "Administration and Commerce" have higher income when compared to individuals in other fields of vocational education. Self-employment of the parents is the most common for the individuals who choose "Natural Resources".

There is also an indication of some regional differences in the schooling choices. Close to 50% of individuals who choose "Natural resources" come from rural areas (versus 21.5% of those who choose high school). On the other hand, of those who choose high school, 60% come from urban areas and similarly for "Administration and Commerce".

Finally, we can see that there are clear differences in the alternative specific characteristics between the schooling choices as well as between men and women. Those who have a high school diploma have a higher mean income and a lower unemployment rate. Men have on average a higher income than women in all the schooling choices, and their unemployment rate tends to be lower.

5 Empirical Analysis

We are interested in modelling the schooling decisions of young people who have just completed their compulsory education. We model the decision the individuals face as a utility-maximizing discrete choice among the different alternatives, where the alternatives are the options faced by a 16-year old leaving compulsory schooling: high school and 7 fields of vocational education. We model the choice using McFadden's conditional logit model. To allow for possible heterogeneity with respect to the various characteristics, we employ a mixed (random coefficients) logit specification. We always perform our estimations separately for boys and girls. We include in our baseline estimations only the 46 per cent sub-sample for whom we have the information on all the regressors.

5.1 Specification

We model the schooling decision as a function of characteristics that vary by alternatives and the choice situation (alternative-specific) as well as characteristics of the individuals (case-specific). We also include an alternative-specific fixed effect (for unobservable characteristics). The conditional logit model specifies that the probability that individual i chooses alternative j conditional on alternative-specific regressors x_{ij} and case-specific regressors z_i is

$$p_{ij} = \frac{\exp(x'_{ij} \beta + z'_i \gamma_j)}{\sum_{l=1}^m \exp(x'_{il} \beta + z'_i \gamma_l)}, j = 1, \dots, m$$

The alternative-specific variables include a variable describing the supply of that alternative and two dummy variables indicating whether the mother and the father have a degree from that field. We use the following variables to describe the supply of upper-secondary education: One, we measure the distance to the closest municipality offering each alternative. Alternatively (unreported), we include a dummy variable indicating the presence of each type of education in the individual's municipality of residence. Finally, we include a variable describing the availability of places for new students for each schooling alternative in the sub-region by calculating the number of students enrolled relative to the size of the cohort (work in progress).

The estimated effects of the supply factors are based on the assumption that the supply of education is exogenous to the individuals' choices. One way in which there may be a possible selection problem is the selection of parents to municipalities based on their occupation. We are,

however, able to control for the parents' income, socioeconomic status and their education (the field and the level of their education) and thus significantly diminish the problem.

We include a number of variables that describe the individual. First, to examine how previous school performance affects choices we use the grade point average in comprehensive school as a measure of academic ability. We also include the average of the grade point average of the students from the same comprehensive school as well as the size of the comprehensive school. Second, we have variables describing family background. We include the level of education and the socioeconomic status of both the mother and the father and the family income. Third, we have variables describing the region where the individual lives. We include the youth unemployment rate in the municipality, a variable indicating whether the region is rural, suburban, or urban, as well as a dummy variable for each province. Finally, we include the nationality and native language of the individual.

In the mixed logit specification we allow individual specific coefficients for the supply variable as well as for the alternative specific constant²⁰.

5.2 Results

5.2.1 Supply of education

Table 6 shows the estimated coefficients for the supply of education (measured by the distance) from the alternative-specific conditional logit and mixed logit specifications. We can see that distance to the nearest municipality offering each field has a negative significant effect on the decision to choose that alternative for both boys and girls. The estimated effects are larger for boys. The results from the mixed logit specification provide evidence that there is some heterogeneity in the effects of supply for boys, but not for girls. For boys, the estimated random coefficient of supply is larger and again statistically significant and the standard deviation of random coefficient is also statistically significant. For girls however, the estimated random coefficient of supply is unchanged and the standard deviation of random coefficient is statistically insignificant.

[Table 6 here]

In Table 7 we report the marginal effects of increasing the distance to a given alternative by 100 kilometres from the alternative-specific conditional logit. As expected most of the own-effects are negative and the cross-effects positive. The estimated marginal effects are relatively small, but statistically significant. Based on the observed cross-effects, for boys particularly "High school" and "Technology and Transport" appear to be close substitutes and for girls the substitution pattern is the strongest for "High School" and "Social and Health Care Services".

[Table 7 here]

²⁰ We have also explored random coefficient for the parents field of education. No significant heterogeneity was visible and the main results remain.

5.2.2 *Parental background and previous school performance*

In previous literature it has been found that parental background characteristics and previous school performance are important factors in determining the schooling decisions. We are able to examine in detail the influence of various parental characteristics as well as the individual's school performance in comprehensive school. We find similar results, but our detailed analysis shows interesting differences in how these factors affect the choice of various alternatives, in particular that different factors are relevant for different alternatives. Table 8 shows the estimates of the marginal effects at the mean from the alternative specific conditional logit estimations for boys (for girls unreported).

[Table 8 here]

From Table 8, we can see that the choice among alternatives is positively affected by having a parent with the same field of education. One exception is that when the mother's education is from the field of "Culture", the son is less likely to choose that field. Similarly, the cross-effects of parents' fields of education (i.e. the effects on the probability of choosing a field that is different from the parents' field) are as expected, mostly negative and significant. The effect of father's education is stronger than that of mother's. For girls, the mother's field of education has no statistically significant effects and the father's field has fewer significant effects (of the expected sign) than for boys.

Having a mother with a degree higher than upper secondary education increases the boys' probability to apply to high school and decreases their probability of choosing "Technology and Transport" and "Administration and Commerce". Father's level of education has the same effect with similar magnitudes. Furthermore, the father's higher level of education also decreases boys' probability to choose "Hotel, Catering and Home Economics". The estimated effects are again smaller for girls. However, the negative effect of both mother's and father's higher education on choosing "Social and Health Care Services" is statistically significant for girls.

In terms of the socioeconomic status of the parents, the main finding that emerges is that the sons of self-employed parents are more likely to choose natural resources. For girls, this result is not statistically significant. For girls, having a mother whose socioeconomic status is "Upper level employee" increases the probability of choosing high school.

Concerning previous school performance, we find that higher average grades in comprehensive school are positively correlated with choosing high school (statistically significant only for boys). Negative correlations between school performance and choices emerge for all vocational fields (for both boys and girls). In particular, the negative effect is large for the field of "Technology and transport" for boys and for the field of "Social and Health Care Services" for girls.

6 Conclusions

We have put together a very detailed and rich individual level dataset on four full cohorts of Finns leaving comprehensive school. We observe their personal characteristics, school performance, information on the comprehensive school they attended, and parents' information. We also observe their location of residence at the time of finishing comprehensive school as well as at the end of the year. The majority of the individuals apply to upper-secondary education and we observe their application and enrolment decisions into vocational schools and high schools. The application data is interesting. It tells us up to 5 choices ranked in order of preference, which together with acceptance and enrolment data provides interesting avenues for research. We also construct measures of supply of education in each municipality and link the municipalities to inter-municipality distances to get measures of distances to municipalities offering each education alternative. Our data also includes regional youth unemployment and alternative-specific characteristics, i.e. mean wages and unemployment rates for men and women.

In this paper, we use this data and employ a conditional logit model to examine the schooling choices of young people and pay particular attention to the effects of educational supply. Our results show that the supply of education, in particular the availability and proximity of education alternatives, plays a statistically significant role in influencing schooling decisions and thus influences the sorting of individuals across the schooling alternatives, at least to some extent. The effect is larger for boys than for girls, and for boys we also find significant heterogeneity of the effect in our random coefficients. Also, in line with prior evidence, we find that parental factors and previous school performance are important determinants of schooling choices. Overall, our analysis shows that boys are very different in their schooling choices and in how they respond to the various determinants.

In work-in-progress, we explore the heterogeneity of the effects and substitution patterns in more detail. We examine the determinants of schooling choices separately for different groups of individuals (e.g. according to parents' income, individual's school performance). We also further explore random coefficients for some of the other determinants and explore the marginal effects at various values of the observed covariates. We also examine the substitution patterns between alternatives in more detail by utilizing the information on the application requests (up to five requests and their ranking). This data directly reveals further information about individuals' preferences over the alternatives.

In another ongoing work, we complement this study by examining how the supply of education affects social exclusion. We study in detail the effect on the probability of continuing in the 9th or 10th grade and of applying and enrolling in upper secondary education, as well as on the drop-out behaviour. Preliminary results indicate that the presence of some particular fields (eg. Technology and Transport) may be important in affecting the decision to participate.

The careful modelling of the schooling choices of individuals forms the foundation for various other studies that examine the progress of individuals through their schooling process, and their schooling outcomes.

References

- Arcidiacono, P. (2004) "Ability sorting and the return to college major", *Journal of Econometrics* 121, 343-375.
- Beffy, M., Fougère, D. and Maurel, A. (2009) "Choosing the field of study in post-secondary education: do expected earnings matter?" IZA Discussion Paper No. 4127.
- Bound, J. and Turner, S. (2007) "Cohort crowding: How resources affect collegiate attainment", *Journal of Public Economics* 91, 877-899.
- Card, D. (2001) "Estimating the return to schooling: progress of some persistent econometric problems", *Econometrica* 69:5, 1127-1160.
- Dustmann, C. (2004) "Parental background, secondary school track choice, and wages", *Oxford Economic Papers* 56, 209-230.
- Frenette, M. (2004) "Access to college and university: does distance to school matter?", *Canadian Public Policy* 30:4, 427-443.
- Gibbons, S. and Vignoles, A. (2009) "Access, choice, and participation in higher education", CEE Discussion Paper No. 101.
- Kaukonen, R. (2009) "Valtaosa uusista ylioppilaista viettää välivuoden opinnoista", *Hyvinvointikatsaus* 1/2009. Tilastokeskus.
- Spiess, C.K. and Wrohlich, K. (2008) "Does distance determine who attends a university in Germany?", IZA Discussion Paper No. 3615.
- Press release of the Ministry of Education (2008) "Sivistyspoliittinen ministeriryhmä linjasi koulutuksen alueelliset tavoitteet vuodelle 2012." 16.5.2008.

APPENDIX A Fields and sub-fields of vocational education (the classification by the Finnish National Board of Education in 1995)

FIELDS OF VOCATIONAL EDUCATION	SUB-FIELDS
Natural Resources Sector	Agriculture Horticulture Fishery Other primary industries Forestry
Technology and Transport Sector	Graphics technology Heating and ventilation Machinery and metal technology Vehicles and transportation Textiles and clothing Food industry Electrical engineering Land survey technology Construction technology Wood industry Surface treatment Paper and chemical industry Seafaring Other technology and transportation
Administration and Commerce Sector	Business and administration
Hotel, Catering and Home Economics Sector	Hotel, restaurant and catering Home economics and cleaning services
Social and Health Care Services Sector	Social and health care services Beauty care
Culture Sector	Crafts and design Communications and visual arts Music Theatre and dance
Humanist and Teaching Sector	Leisure activities Physical education

Table 1. Regional availability of the schooling alternatives

Schooling alternative	Distance to the nearest municipality offering each alternative (km)				Share of municipalities with each alternative (%)	Transition probability (%)		No within variation	
	Mean	Max	Std. Dev. (between)	Std. Dev. (within)		From 0 to 0	From 0 to 1	Alt. offered 00-03	Alt. not offered 00-03
High school	8.1	74	10.5	0.0	68.3%	0.0	0.0	100.0	100.0
Natural Resources	33.3	144	22.4	4.7	17.3%	4.5	0.6	91.4	98.6
Technology and Transport	24.8	320	25.6	2.5	27.8%	2.4	0.8	96.1	97.8
Administration and Commerce	29.7	144	22.1	6.2	19.5%	8.9	1.7	85.8	95.6
Hotel, Catering and Home Economics	24.9	125	20.1	5.5	27.8%	7.1	2.4	90.2	95.0
Social and Health Care Services	38.1	434	38.9	5.0	14.8%	6.0	1.1	89.7	98.6
Culture	39	434	38.0	17.3	13.8%	8.7	2.6	82.2	95.8
Humanist and Teaching	63.6	570	49.1	15.9	6.8%	9.2	0.9	85.3	98.2
Any vocational field	17.1	125	17.9	2.7	43.0%	2.5	2.5	94.8	95.2
Any schooling alternative	7.7	74	10.3	0.6	70.3%	0.1	0.8	99.4	99.0

Note. If the schooling alternative offered in the municipality, distance defined to be 2km.

Table 2. Schooling choices by 1st request

Schooling choice	Boys	Girls	All
High school	49.9	68.8	59.1
Natural Resources	2.1	1.6	1.9
Technology and Transport	35.2	2.8	19.4
Administration and Commerce	7.1	4.4	5.7
Hotel, Catering and Home Economics	3.7	8.0	5.8
Social and Health Care Services	0.5	10.9	5.6
Culture	1.4	2.7	2.0
Humanist and Teaching	0.2	1.0	0.6

Table 3. Proximity of schooling alternatives conditional on schooling choices

Schooling alternative	Distance to the nearest municipality offering an alternative (km)							
	high school	natural resources	technology and transport	administration and commerce	hotel, catering and home economics	social and health services	culture	humanist and teaching
High school	3,1	23,0	10,0	13,4	10,3	17,5	20,7	46,8
Natural Resources	4,3	28,2	18,9	25,3	19,2	32,6	36,9	62,7
Technology and Transport	3,5	24,6	11,4	16,6	12,2	20,8	24,2	53,0
Administration and Commerce	3,1	24,3	9,3	11,8	9,5	17,0	21,2	48,3
Hotel, Catering and Home Economics	3,2	24,6	10,6	15,3	10,9	20,1	24,7	53,7
Social and Health Care Services	3,3	25,0	10,9	16,1	11,6	19,4	23,7	52,4
Culture	3,2	24,5	12,5	15,6	12,4	19,3	18,9	51,6
Humanist and Teaching	3,4	24,9	12,4	16,8	12,6	21,9	25,4	37,0
Outside option (not to apply)	2,9	23,7	9,5	12,8	9,7	17,3	19,6	43,0
Total	3,2	23,7	10,6	14,4	10,9	18,7	22,1	49,0

Table 4. Decisions of peers from the comprehensive school

Schooling alternative	Share of individuals from the comprehensive school who have chosen the given schooling alternative (%)							
	high school	natural resources	technology and transport	administration and commerce	hotel, catering and home economics	social and health services	culture	humanist and teaching
High school	63,0	1,5	16,8	5,4	5,0	5,1	1,8	0,5
Natural Resources	52,2	7,2	20,2	5,3	5,9	5,7	2,0	0,7
Technology and Transport	54,1	1,9	23,7	5,5	5,6	5,7	2,0	0,6
Administration and Commerce	55,9	1,6	17,7	10,5	5,4	5,5	1,9	0,5
Hotel, Catering and Home Economics	54,5	1,9	19,2	5,7	9,8	5,5	2,0	0,5
Social and Health Care Services	55,0	1,8	19,1	5,7	5,4	9,5	2,0	0,5
Culture	55,0	1,7	18,6	5,5	5,4	5,6	6,6	0,6
Humanist and Teaching	54,4	2,1	19,5	5,7	5,2	5,3	2,2	4,8
Total	59,6	1,7	18,5	5,7	5,4	5,5	2,0	0,5

Note. The observations from comprehensive schools with less than 10 individuals in a given year are excluded.

Table 5. Conditional characteristics of individuals, parents, regions, and alternatives

Schooling alternative		High school	Natural Resources	Technology and Transport	Administration and Commerce	Hotel, Catering and Home Economics	Social and Health Care Services	Culture	Humanist and Teaching	All
Individual characteristics										
Share of men (%)		43.2	58.5	93.0	62.9	32.7	4.9	35.1	16.3	51.2
Average grade of theoretical subject		8.2	6.8	6.5	7.0	6.8	6.9	7.0	6.9	7.6
Average grade of all subjects		8.3	7.1	6.9	7.2	7.1	7.3	7.4	7.3	7.8
Average number of pupils in comprehensive school		33.5	28.1	31.3	33.6	31.0	32.8	32.1	31.9	32.6
Parental characteristics (mother father)										
Field of education (%)										
General education	mother	4.4	2.4	2.9	4.4	3.0	3.7	3.7	4.2	4.0
	father	3.8	1.1	1.7	2.6	2.5	2.5	1.8	0.5	3.2
Natural Resources	mother	4.2	8.8	3.3	2.7	3.1	2.6	3.8	3.8	3.9
	father	12.0	34.9	12.1	8.4	11.4	10.6	11.8	12.2	12.1
Technology and Transport	mother	7.7	11.0	12.5	11.3	13.3	12.3	10.4	13.1	9.3
	father	53.8	52.7	73.4	67.6	68.3	70.3	65.7	69.0	59.1
Administration and Commerce	mother	28.1	21.0	24.0	28.0	23.5	23.7	24.4	21.2	26.7
	father	15.2	5.5	6.6	12.2	7.9	7.6	10.7	6.6	12.8
Hotel, Catering and Home Economics	mother	11.6	26.3	24.3	19.6	24.7	25.3	18.7	20.8	15.8
	father	1.2	1.3	1.1	1.8	2.5	1.8	0.9	2.0	1.3
Social and Health Care Services	mother	30.3	25.1	28.1	28.3	27.7	28.6	29.8	30.0	29.5
	father	4.8	0.9	1.5	2.5	2.4	3.0	3.4	3.0	3.9
Culture	mother	1.9	1.6	1.5	1.6	1.4	0.9	3.3	1.5	1.8
	father	1.0	0.6	0.4	0.7	0.7	0.8	0.6	1.0	0.8
Humanist and Teaching	mother	11.7	3.7	3.2	4.0	3.3	2.9	5.7	5.4	8.8
	father	4.7	0.5	1.0	1.8	1.2	0.9	2.3	2.5	3.5
Field of education unobserved	mother	12.4	22.1	25.4	22.2	29.4	27.2	19.7	19.0	17.5
	father	18.0	34.4	32.7	29.2	33.7	33.1	27.4	26.4	23.1
Level of education (%)										
Secondary education	mother	43.0	71.4	71.4	65.1	72.3	73.3	63.3	67.3	52.7
	father	44.2	80.8	76.6	66.9	78.5	78.0	65.9	73.6	54.4
College- level	mother	33.9	22.2	22.8	26.8	21.3	21.9	27.4	25.0	30.2
	father	22.0	13.6	15.4	20.7	13.2	13.7	19.3	17.3	19.9
Higher education	mother	23.1	6.3	5.8	8.1	6.4	4.8	9.3	7.7	17.2
	father	33.8	5.6	8.0	12.4	8.3	8.2	14.8	9.1	25.7
Socioeconomic status (%)										
Self-employed persons	mother	9.5	26.5	10.2	9.0	8.9	9.8	10.8	9.3	9.9
	father	17.0	38.3	18.4	17.4	16.9	17.2	20.2	14.2	17.7
Upper-level employees	mother	22.1	5.3	5.8	9.2	5.7	5.1	8.8	8.4	16.0
	father	27.6	4.6	6.7	11.5	7.5	6.8	12.6	7.8	20.2
Lower-level employees	mother	42.1	27.8	35.2	39.7	32.5	34.7	38.1	37.1	39.4
	father	19.8	8.0	13.6	17.7	13.3	14.7	14.8	16.4	17.7
Manual workers	mother	14.0	22.5	28.1	23.9	29.7	28.8	22.2	25.2	19.1
	father	25.8	33.1	46.0	40.2	45.1	45.7	40.8	50.0	32.6
Students	mother	1.9	2.1	2.2	2.4	2.9	2.5	2.4	3.4	2.1
	father	0.7	0.8	0.7	0.9	0.5	1.0	0.7	0.0	0.7
Pensioners	mother	1.7	3.0	2.8	2.3	2.6	2.8	2.5	3.4	2.1
	father	3.5	5.8	5.1	4.7	5.1	5.1	3.8	3.7	4.1
Others	mother	8.7	12.7	15.7	13.6	17.7	16.4	15.2	13.1	11.4
	father	5.5	9.4	9.5	7.6	11.6	9.5	7.2	7.8	7.0
Information on a parent missing (%)	mother	10.9	17.6	18.5	18.4	23.4	22.4	18.7	18.3	14.7
	father	16.5	15.5	23.3	23.6	31.1	26.8	25.5	22.1	19.8
Family income (€)		54,334	34,655	40,254	44,128	39,366	39,759	43,063	40,925	54,334
Regional characteristics										
Urbanization of municipality (%)										
Rural		21.5	48.7	27.3	20.0	23.7	25.7	26.2	30.3	23.5
Densely populated		18.7	22.8	22.2	20.6	21.3	22.4	22.8	19.3	19.9
Urban		59.8	28.5	50.6	59.3	55.0	52.0	51.1	50.4	56.5
Youth unemployment rate (%)		15.3	18.1	16.8	15.7	16.5	16.5	16.8	16.9	15.8
Alternative specific characteristics										
Mean income of male (€)		38,177	12,453	23,256	24,610	18,398	21,845	16,571	22,426	22,217
Mean income of female (€)		24,228	10,571	14,744	18,036	14,232	17,731	11,994	17,040	16,072
Unemployment rate of male (%)		5.3	9.4	11.6	10.1	14.6	6.4	18.3	6.9	10.3
Unemployment rate of female (%)		4.9	14.3	16.0	10.4	13.4	6.2	18.7	11.5	11.9

Table 6. Estimated coefficients for supply of education (distance to the nearest municipality, km)

Distance to the nearest municipality	Boys		Girls	
	Asclgit	Mixed logit	Asclgit	Mixed logit
Mean	-0.010*** (0.001)	-0.014*** (0.002)	-0.005*** (0.001)	-0.005*** (0.001)
Std.E	-	0.011*** (0.003)	-	0.003 (0.003)

Number of obs 132,984 132,984 123,408 123,408
 Log likelihood -11812.06 -11802.61 -9797.64 -9795.40

Notes: Standard errors in parentheses, *** significant at 0.1 per cent

Estimations contain 16,623 boys and 15,426 girls.

All estimations Include mother's and father's socioeconomic status, level and field of education, family income, mother tongue, individual's average grade in comprehensive school, peer's average grade in comprehensive school, size of comprehensive school, urbanization of the municipality and youth unemployment

Table 7. Marginal effects of supply from the conditional logit specification (100km)

Boys	High school	Natural Resources	Technology and Transport	Administration and Commerce	Hotel, Catering and Home Economics	Social and Health Care Services	Culture	Humanist and Teaching
					Hotel, Catering and Home Economics	Social and Health Care Services		
High school	-0.193*** (0.037)	0.004*** (0.001)	0.125*** (0.024)	0.044*** (0.008)	0.011*** (0.002)	0.001 (0.056)	0.008*** (0.002)	-
Natural Resources	0.004*** (0.001)	-0.006*** (0.001)	0.001*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000 (0.000)	0.000*** (0.000)	-
Technology and Transport	0.125*** (0.024)	0.001*** (0.000)	-0.139*** (0.019)	0.009*** (0.002)	0.002*** (0.001)	0.000 (0.012)	0.002*** (0.000)	-
Administration and Commerce	0.044*** (0.008)	0.000*** (0.000)	0.009*** (0.002)	-0.055*** (0.008)	0.001*** (0.000)	0.000 (0.004)	0.001*** (0.000)	-
Hotel, Catering and Home Economics	0.011*** (0.002)	0.000*** (0.000)	0.002*** (0.001)	0.001*** (0.000)	-0.015*** (0.003)	0.000 (0.001)	0.000*** (0.000)	-
Social and Health Care Services	0.001 (0.056)	0.000 (0.000)	0.000 (0.012)	0.000 (0.004)	0.000 (0.001)	-0.001 (0.074)	0.000 (0.001)	-
Culture	0.008*** (0.002)	0.000*** (0.000)	0.002*** (0.000)	0.001*** (0.000)	0.000*** (0.000)	0.000 (0.001)	0.010*** (0.002)	-
Humanist and Teaching	0.000 (0.007)	0.000 (0.000)	0.000 (0.001)	0.000 (0.001)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	-
Pr(choice = y l selected)	0.753	0.005	0.160	0.056	0.014	0.001	0.010	0.000
Girls								
High school	-0.036 (0.028)	0.002 (0.013)	0.003** (0.001)	0.007*** (0.002)	0.006*** (0.002)	0.012*** (0.003)	0.005 (0.028)	0.001 (0.009)
Natural Resources	0.002 (0.013)	-0.002 (0.014)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Technology and Transport	0.003** (0.001)	0.000 (0.000)	-0.003*** (0.001)	0.000** (0.000)	0.000** (0.000)	0.000** (0.000)	0.000 (0.000)	0.000 (0.000)
Administration and Commerce	0.007*** (0.002)	0.000 (0.000)	0.000** (0.000)	-0.007*** (0.002)	0.000** (0.000)	0.000*** (0.000)	0.000 (0.000)	0.000 (0.000)
Hotel, Catering and Home Economics	0.006*** (0.002)	0.000 (0.000)	0.000** (0.000)	0.000** (0.000)	-0.007*** (0.002)	0.000*** (0.000)	0.000 (0.000)	0.000 (0.000)
Social and Health Care Services	0.012*** (0.003)	0.000 (0.000)	0.000** (0.000)	0.000*** (0.000)	0.000*** (0.000)	-0.013*** (0.003)	0.000 (0.001)	0.000 (0.000)
Culture	0.005 (0.028)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.001)	-0.005 (0.030)	0.000 (0.000)
Humanist and Teaching	0.001 (0.009)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	-0.001 (0.009)
Pr(choice = y l selected)	0.918	0.004	0.006	0.015	0.015	0.028	0.011	0.002

Notes: Standard errors in parentheses, ** significant at 1 per cent, *** significant at 0.1 per cent

Table 8. Marginal effects for boys from the conditional logit specification

	High school	Natural Resources	Technology and Transport	Admin and Commerce	Hotel, Catering and Home Economics	Social and Health Care Services	Culture	Humanist and Teaching
Mother's field of education								
General/high school	0.024** (0.008)	-0.001** (0.000)	-0.015** (0.005)	-0.005** (0.002)	-0.001** (0.000)	-0.000 (0.007)	-0.001** (0.000)	0.000 (0.000)
Natural Resources	-0.001* (0.000)	0.001** (0.000)	-0.000* (0.000)	-0.000* (0.000)	-0.000* (0.000)	-0.000 (0.000)	-0.000** (0.000)	0.000 (0.000)
Technology and Transport	-0.016** (0.006)	-0.000** (0.000)	0.018** (0.006)	-0.001** (0.000)	-0.000** (0.000)	-0.000 (0.002)	-0.000** (0.000)	0.000 (0.000)
Administration and Commerce	-0.006** (0.002)	-0.000** (0.000)	-0.001** (0.000)	0.007** (0.002)	-0.000** (0.000)	-0.000 (0.001)	-0.000** (0.000)	0.000 (0.000)
Hotel, Catering and Home Economics	-0.001** (0.000)	-0.000* (0.000)	-0.000** (0.000)	-0.000** (0.000)	0.002** (0.000)	-0.000 (0.000)	-0.000** (0.000)	0.000 (0.000)
Social and Health Care Services	-0.000 (0.007)	-0.000 (0.000)	-0.000 (0.002)	-0.000 (0.001)	-0.000 (0.000)	0.000 (0.010)	0.000 (0.000)	0.000 (0.000)
Culture	-0.001** (0.000)	-0.000* (0.000)	-0.000** (0.000)	-0.000** (0.000)	-0.000** (0.000)	-0.000 (0.000)	-0.001** (0.000)	0.000 (0.000)
Humanist and Teaching	-0.000 (0.001)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Father's field of education								
General/high school	0.071*** (0.013)	-0.002*** (0.000)	-0.046*** (0.007)	-0.016*** (0.003)	-0.004*** (0.001)	-0.000 (0.020)	-0.003*** (0.001)	-0.000 (0.002)
Natural Resources	-0.002*** (0.000)	0.003*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000 (0.000)	-0.000*** (0.000)	-0.000 (0.000)
Technology and Transport	-0.049*** (0.008)	-0.000*** (0.000)	0.054*** (0.005)	-0.004*** (0.001)	-0.001*** (0.000)	-0.000 (0.005)	-0.001*** (0.000)	-0.000 (0.001)
Administration and Commerce	-0.020*** (0.003)	-0.000*** (0.000)	-0.004*** (0.001)	0.025*** (0.003)	-0.000*** (0.000)	-0.000 (0.002)	-0.000*** (0.000)	-0.000 (0.000)
Hotel, Catering and Home Economics	-0.005*** (0.001)	-0.000*** (0.000)	-0.001*** (0.000)	-0.000*** (0.000)	0.007*** (0.001)	-0.000 (0.001)	-0.000*** (0.000)	-0.000 (0.000)
Social and Health Care Services	-0.000 (0.027)	-0.000 (0.000)	-0.000 (0.006)	-0.000 (0.002)	-0.000 (0.001)	0.000 (0.035)	-0.000 (0.000)	-0.000 (0.000)
Culture	-0.004*** (0.001)	-0.000*** (0.000)	-0.001*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000 (0.000)	0.005*** (0.001)	-0.000 (0.000)
Humanist and Teaching	-0.000 (0.003)	-0.000 (0.000)	-0.000 (0.001)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.004)
Mother's level of education baseline: upper secondary education								
College	0.052*** (0.011)	0.000 (0.001)	-0.031*** (0.008)	-0.016*** (0.005)	-0.003 (0.002)	-0.000 (0.009)	-0.002 (0.002)	-0.000 (0.004)
Higher degree	0.079*** (0.015)	-0.000 (0.001)	-0.055*** (0.012)	-0.023*** (0.006)	0.001 (0.003)	-0.000 (0.001)	-0.002 (0.003)	-0.000 (0.003)
Father's level of education baseline: upper secondary education								
College	0.051+ (0.026)	-0.000 (0.001)	-0.032 (0.011)	-0.011* (0.005)	-0.005* (0.002)	-0.001 (0.035)	-0.002 (0.002)	0.000 (0.002)
Higher degree	0.080** (0.024)	-0.002 (0.001)	-0.050*** (0.014)	-0.021** (0.007)	-0.008** (0.003)	-0.001 (0.031)	0.001 (0.003)	0.000 (0.008)
Mother's socioeconomic status baseline: self-employed								
Upper-level employees	0.016 (0.049)	-0.004** (0.001)	-0.029 (0.018)	0.016 (0.011)	0.002 (0.004)	-0.001 (0.058)	-0.001 (0.003)	0.000 (0.013)
Lower-level employees	0.026 (0.046)	-0.006*** (0.001)	-0.025 (0.016)	0.004 (0.008)	0.003 (0.003)	-0.001 (0.059)	-0.000 (0.003)	0.000 (0.005)
Manual workers	-0.044* (0.021)	-0.002* (0.001)	0.028+ (0.015)	0.014 (0.009)	0.007 (0.005)	-0.000 (0.002)	-0.002 (0.002)	0.000 (0.006)
Students and pensioners	-0.008 (0.032)	-0.002 (0.001)	-0.007 (0.020)	0.012 (0.014)	0.002 (0.006)	-0.000 (0.014)	0.002 (0.005)	-0.000 (0.013)
Other	-0.021 (0.024)	-0.003*** (0.001)	0.009 (0.017)	0.010 (0.011)	0.001 (0.004)	-0.000 (0.002)	0.004 (0.005)	0.000 (0.001)
Information missing	0.037 (0.067)	-0.004** (0.001)	-0.044+ (0.027)	0.013 (0.020)	-0.005 (0.005)	-0.002 (0.073)	0.005 (0.008)	-0.000 (0.010)
Father's socioeconomic status baseline: self-employed								
Upper-level employees	0.002 (0.039)	-0.003** (0.001)	-0.023 (0.014)	0.014 (0.009)	0.009+ (0.005)	0.001 (0.045)	-0.000 (0.003)	0.000 (0.010)
Lower-level employees	0.017 (0.055)	-0.005*** (0.001)	-0.020 (0.014)	0.004 (0.008)	0.001 (0.003)	0.001 (0.049)	0.001 (0.003)	0.000 (0.048)
Manual workers	-0.030 (0.056)	-0.001 (0.001)	0.016 (0.018)	0.009 (0.009)	0.004 (0.003)	0.001 (0.074)	0.000 (0.003)	0.000 (0.013)
Students and pensioners	0.014 (0.062)	-0.001 (0.001)	-0.021 (0.020)	0.010 (0.014)	0.000 (0.005)	0.001 (0.071)	-0.003 (0.004)	-0.000 (0.013)
Other	0.020 (0.128)	-0.002+ (0.001)	-0.011 (0.029)	-0.012 (0.011)	0.006 (0.007)	0.003 (0.160)	-0.003 (0.003)	-0.000 (0.012)
Information missing	0.008 (0.054)	-0.005*** (0.001)	-0.025 (0.025)	0.016 (0.018)	0.009 (0.009)	-0.002 (0.053)	-0.001 (0.005)	-0.000 (0.010)
Average grade of individual	0.387***	-0.009***	-0.262***	-0.074***	-0.025***	-0.002	-0.014***	-0.000 (.)

Notes: Standard errors in parentheses, + significant at 10 per cent, * significant at 5 per cent, ** significant at 1 per cent, *** significant at 0.1 per cent