Preferences for lifetime earnings, earnings risk and nonpecuniary attributes in choice of higher education*

Lars J. Kirkebøen†

March 24, 2012

Abstract

Expected earnings are reasonably considered to influence individuals' choice of education. However, the presence of nonpecuniary attributes and the different choice set available to prospective students make identification of this relationship difficult. This paper employs a conditional logit model on exceptionally rich application data, which are likely to reflect the actual preferences of the applicants, given their individual choice sets. Controlling for several nonpecuniary attributes, average lifetime earnings is shown to strongly influence educational choice. A one-percent earnings increase increases the number of male applicants to an education by about 6 percent and female applicants by about 3 percent. However, other attributes also matter, in particular earnings risk. Increasing both earnings and risk as they correlate in the cross section, there is little if any effect on choices.

Keywords: Rank-ordered logit, nested logit, field of study

JEL Classification: J24, J31, C25

1 Introduction

There is a long tradition in economics for studying how expected earnings influence choice of education, e.g. Boskin (1974); Berger (1988); Arcidiacono et al (2011); Beffy *et al* (2011).

*I am grateful to Rolf Aaberge, John Dagsvik, Oddbjørn Raaum and Terje Skjerpen for helpful discussions and comments. The project is funded by the Ministries of Education and Research, Labour, Trade and Industry and Health and Care Services, who have also contributed helpful comments.

†Research Department, Statistics Norway. E-mail: kir@ssb.no

This question is of great relevance to understand the functioning of the labor market in general as well as to specific policy questions. For example, large and persistent earnings differences exist between fields of education, which in turn influence strongly on gender differences in earnings. Do prospective students respond to this, potentially increasing the supply of fields in high demand, and eventually earnings gaps? Also, if there is a need for an increase in the supply with a given education, say, of health care professional as the population ages, it is relevant to know what earnings increase can provide such an increase in students.

However, this literature is still relatively small. One reason may be the problems involved in estimating this relationship. The educational alternatives will also have nonpecuniary attributes, which may both influence choices and be correlated with observed earnings. E.g. Zafar (2009); Arcidiacono et al (2011); Beffy *et al* (2011); Wiswall and Zafar (2011) find differences in average preferences for fields. Thus, failure to control for these must be expected to give an omitted-variable bias in the estimated significance of earnings for educational choice. Also, comparative advantage and even choice sets vary between prospective students,[1] such that even to the extent that individuals do maximize expected earnings, they will do so subject to constraints that are generally hard to identify for the researcher.

This paper specifies a simple model for formation of earnings expectations and choice of education. The model is then estimated, using rich Norwegian register data. Of particular relevance is the fact that the Norwegian system for admission to higher education is very centralised, and, for the most part, strictly meritocratic. Thus, as the admission process uses well-defined data, with essentially no discretion, it is possible to make reasonable assumptions of the applicants' knowledge of their opportunities at the time of application. Furthermore, the admission makes it likely that students state their true preferences, conditional on their (perceived) opportunities, and rank several alternatives, thus increasing the information content of the data. Finally, earnings, earnings risk and several nonpecuniary attributes are all controlled for in the estimations, thus both providing a richer picture of the determinants of

[1] Paglin and Rufolo (1990) find evidence that mathematical ability is an important determinant of college field choice, while Arcidiacono et al (2011) find that self-reported relative skill in all fields matter. Nielsen and Vissing-Jorgensen (2005) argues that it is relevant to control for choice sets in the study of educational choice, while Desposato (2005) argues that choice set selection in general may have a large impact on conditional logit estimates.

educational choice, and reducing the scope for omitted-variable bias in the estimated effect of earnings.

The paper specifies a nested logit model, where the utility associated with each educational alternative (nest) depends on a set of observable attributes of the alternatives. Choice of courses within nests depend on an unobserved random term. Thus the model simplifies to a conditional logit model, where the number of available courses for each alternative enter the modeled utilities. A rank-ordered logit model is estimated to fully use the information content in the data.

Most prior studies use either the actual occupation or completed education, i.e. the final outcome of the total process initiated with the application (e.g. Boskin (1974); Berger (1988); Beffy *et al* (2011)) or data from surveys with relatively few observations (e.g. Arcidiacono et al (2011); Zafar (2009); Wiswall and Zafar (2011)). The Norwegian application data have the combined virtues of having a large number of observations and being highly relevant for the study of educational choice. Norway has a very centralized system for applications to higher education. Thus, almost all applications for almost every kind of higher education is captured by the application data, giving a sample size of about 50,000 individuals per year for the years 2004-2009. Furthermore, applicants state not only their most-preferred course, but rank up to ten courses, further increasing the amount of information. Also, the data are likely to express actual preferences. As opposed to survey data, the application data is high-stakes. Furthermore, the possibility to rank up to ten reduces the applicants' need to apply strategically, as they are very likely to have one of their wishes granted. Still, the applicants can be expected to take into consideration the probability of admission. However, as the admission process is almost entirely mechanical and approximately the same data is available to the researcher and the applicants, the likely perceived choice sets can be reconstructed for the analysis.

Average lifetime earnings is indeed found to matter for the choice of field and level of education. A one-percent increase in earnings for a given field increases the number of applicants by about 6 percent for men and about 3 percent for women. Controlling for earnings risk has a large impact on the estimated effect of earnings. High-earning educational alternatives carry more risk, and the negative effect of the latter partly offsets the positive effect of the former. This is particularly true for women, who are found to have less of a preference for earnings, and to be more deterred by risk.

The estimates are mostly stable over time, and not very sensitive to the choice of earnings measure. However, the specification of the choice set have some influence on the results. The applicants do tend to choose educations similar to their parents, but controlling for this has little impact on the estimated preference for earnings.

The paper proceeds as follows: Section 2 presents related literature, Section 3 the institutional setting, Section 4 the model and data and Section 5 the results from the estimation. Section 6 concludes.

2 Related literature

While choice of field is less investigated than the choice of level of education, the study of how expected earnings influence choice of educational field, major or occupation have a long tradition in economics, dating back at least to Boskin (1974). Boskin (1974) finds that potential earnings explains a part of the difference in occupational choice for all race and gender groups, while Berger (1988) finds that future earnings streams matter more than initial earnings.

Recent findings are mixed. Montmarquette *et al* (2002); Boudarbat (2008) find a clear effect of earnings on choice of major. Estimating a dynamic model of major choice, Arcidiacono (2004) finds a clear preference for earnings, but monetary returns explain little of the sorting across majors. Boudarbat and Montmarquette (2007) find a small effect of earnings on choice of field of study, and no effect for some combinations of gender and parental education. Controlling for a range of nonpecuniary attributes, Zafar (2009) finds no clear effect of subjective earnings expectations, but the few observations give little power. Beffy *et al* (2011) finds a statistically significant, but small effect of earnings. Finally, Arcidiacono et al (2011) find sizeable effects of earnings on major choice.

A crucial point in the estimation of the significance of earnings for educational choice is how earnings expectations are formed. Traditionally, economists have been reluctant to collect or use survey data on subjective expectations. Rather, expectations have been assumed to be rational, with individuals acting on the basis of the same earnings function that the researcher estimates, i.e. earnings depend both on educational choice and other characteristics, such as ability. Examples of studies using this approach are Willis and Rosen (1979); Manski and Wise (1983); Boskin (1974); Berger (1988), and more recently Boudarbat (2008). However, as argued by Manski (1993), the facts that such estimations are complicated and that the approach chosen and results obtained vary between studies suggest that this is not necessarily a realistic description of expectation formation.

One possible alternative suggested by Manski (1993) and used e.g. by Rochat and Demeulemeester (2001) and Boudarbat and Montmarquette (2007) is to simply use average earnings for educational groups, unconditional on other characteristics.

Dominitz and Manski (1996); Betts (1996); Zafar (2011) find that students mostly are able to meaningfully assess expected earnings and earning differences between different educations. Following this, Arcidiacono et al (2011); Zafar (2009); Wiswall and Zafar (2011) have studied educational choice, using data on subjective expectations. Wiswall and Zafar (2011) move one step further and provide information to students, measuring how this influence their assessed probabilities of graduating with a given major. While these studies generally find that respondents largely give meaningful responses when questioned about earnings expectations, and may also revise their expectations in a reasonable way when exposed to more information, the analysis of choices is limited by the small sample sizesAlso, the samples are also selective, typically from one specific selective university, making it difficult to assess the relevance of the findings for other groups of students or potential students. Studies investigating subjective expectations often find that these vary considerably, e.g. Dominitz and Manski (1996); Betts (1996); Zafar (2011), thus motivating the use of such. However, these studies give little guidance on how to best model earnings expectations in the absence of expectations data.

Some studies link educational choices and risk. Flyer (1997) finds that the job-match uncertainty implies an option value, valued by the students. Saks and Shore (2005) find that individuals with higher wealth choose riskier careers, suggesting that risk aversion varies between individuals, and that it matters for educational choices. Nielsen and Vissing-Jorgensen (2005) find that risk, in the transitory and in particular permanent income shocks, impacts negatively on the probability an education is chosen. Also related to risk, Rochat and Demeulemeester (2001) and Montmarquette *et al* (2002) find that a higher chance of completion matter.

Nonpecuniary attributes in general also matter for educational choice. Arcidiacono et al (2011); Beffy *et al* (2011); Nielsen and Vissing-Jorgensen (2005) all find that there are differences in average preference between different fields. Zafar (2009) link choice of major to different attributes of the studies and following careers, finding that nonpecuniary attributes explain a large share of the variation in choices.

Finally, comparative advantages are also found to influence the choice of education. Paglin and Rufolo (1990) emphasize the differences between different types of human capital, i.e. verbal and quantitative skills, and find that comparative advantage accounts for male-female differences in occupational choices. Arcidiacono et al (2011) find that perceived comparative advantages across major contributes to explaining major choice.

3 Institutional setting

Following the Bologna process, higher education in Norway is mostly organized in three-year Bachelor and five-year Masters degree. The higher education sector consists of eight universities[2], eight specialized university institutions and a number of university colleges. The universities provide undergraduate and postgraduate educations in a range of fields, while each specialized university institutions focus on one subject, e.g. business, architecture, veterinary science, sports or theology. Among the university colleges, the most important in terms of number of students are the 22 public regional university colleges. These mostly provide undergraduate professionally oriented courses, such as nursing, teaching, commerce and engineering. The entire sector is dominated by public institutions, with about 85 percent of the students attending one of the universities - which are all public, a public specialized university institu-

[2]Increased from four in 2005 through the conversion of one specialized university institutions and three regional university colleges.

tion or a public university college. The single significant exception is a private business school with about 10 percent of the total number of students. This school also charges a significant tuition fee, which is otherwise absent.

Also, while several of the before-mentioned institutions offer shorter one or two-year courses - which can make up a part of a Bachelor or Masters degree - there is also a number of private institutions providing short vocational and recreational courses, which are not counted as a part of the higher education sector e.g. in official statistics.

Application to higher education is very centralized. Except for the before-mentioned private business school, a single body, Samordna opptak (SO), organizes applications and admissions to all major institutions. Applicants submit a single application to SO, ranking up to ten specific courses, potentially at different universities or university colleges. SO then handles the application process, allocating students to courses according to the number of places at each course and the students' qualifications and admission scores.

In order to qualify for a course an applicant need to qualify for higher education in general, this is mostly achieved by completing the academic track in upper secondary school.[3] Some courses (e.g. science, engineering, medicine) requires specific subjects in math and science from upper secondary school. A few courses have other requirements, e.g. for students entering two-year engineering courses from vocational school, and some arts courses.

Qualified applicants are admitted to courses based on their admission scores, with the students with the higher score getting priority in case of a surplus of applicants. Some of places at some courses are set aside to different quotas (e.g. students from northern parts of Norway at some institutions), however the bulk of the places and applicants are in the two main quotas: improved GPA and unimproved GPA. In the unimproved GPA quota, applicants compete with admission scores calculated as the GPA they got leaving upper secondary school, i.e. average (original) grades and potentially extra points. Grades range from 1 to 6 (only integer values), grade point is calculated as 10 times average grade (with two decimal places). Extra points are awarded for choosing science subjects (max 4 points) or focusing on subjects in upper

[3]There is a range of less common ways to qualify, including higher education, and, as long as some further requirements are met, completed vocational education and for those at least 23 years old any combination of work and schooling for at least five years.

secondary (also max 4 points). There are specific rules for some courses, e.g. 2 extra points for women at some male-dominated courses, and medicine has their own implementation of this quota. Improved GPA includes any changes to the grades as the applicants have redone or taken more secondary school subjects after leaving secondary school,[4] the extra points mentioned above and some more for age, education and military service. Medicine and some other courses have separate regulations for extra points.

While an applicant may know for certain whether he satisfies the formal qualification requirements for a course, he does not know whether he will be admitted. There are two sources of uncertainty: First, the exact score required to be admitted is unknown at the time of application, as this will depend on the number and scores of the other applicants, both unknown by the applicant. However, minimum admission scores for previous years is available from SO, such that the applicants can make an informed guess when applying. Also, applicants still in secondary school in April when applying will not know their final grades and GPA, as these are set in May or June.

4 Model and data

This section presents the empirical model to be estimated, and the data to be used in the estimations. Choices of higher education are made from individual-specific choice sets, and are assumed to depend on expected earnings, nonpecuniary attributes and a random term, which are discussed in turn.

While choice of education is an inherently dynamic process, where choices at one stage influences the options and pay-offs a later stages, a static model of choice of higher education will be estimated. Thus, a limitation of the model is that it does not model earlier educational choices, but rather takes the applicants previous qualifications as given. A richer model could include choices through secondary school. This is outside the scope of this paper.

[4] A number of students spend much time improving their grades to get competitive courses such as medicine.

4.1 Choice of education

There is a total of C different specific courses. These are classified into J different broader educational alternatives (henceforth educations), with each education j consisting of a set C_j of different courses.

As described in Section 3, admission is strongly meritocratic. Thus, each applicant will have an individual-specific choice set, based on her formal qualification, her admission score, the rankings of the other applicants and their admission scores. [5] While there is uncertainty about the two latter at the time of application, these are still exogenous to an individual applicant, such that they can be summarised in an admission score required to qualify. For the time being, I will disregard the uncertainty

Thus, based on formal qualifications, admission score and the required scores of the different courses, an applicant has a choice set C_{ij} of courses within education j that she can be admitted to, containing m_{ij} courses. m_{ij} is smaller or equal to the number of courses in C_j, and may be zero - indicating that at the applicant will not be admitted to any course within this education, and hence can not choose this particular education. The total set of educations j available to the applicant, i.e. with $m_{ij} > 0$ is denoted Ω_i, such that the full choice set of an applicant, including all courses summed across all educations, is $\{c | c \in C_{ij}, j \in \Omega_i\}$. As choice sets are determined by qualifications and admission scores, these will vary between individuals.

An applicant also has preferences for a range of attributes of the courses. Some of these relate to the careers that follow from choosing a career, such as earnings, earnings risk, unemployment and working time. Other relate to the consumption value and cost of studying, and may include e.g. the effort required to follow a particular course, peer students or the location of the institution.

These preferences are revealed through the ranking of courses in the application. The applicant is assumed to evaluate all courses available to her and choose the most attractive one. Thus, the applicant chooses course c, within education j, with the highest utility U_{icj},

[5] This is true for a large majority of the applicants. Those applicants who get discretionary treatment are disregarded in the analysis.

i.e. such that:

$$U_{ijc} = \max_{c' \in C_{ij'}, j' \in \Omega_i} U_{ij'c'} \tag{1}$$

The utility from course c in education j depend on expected earnings, earnings risk and nonpecuniary attributes. In the following a simple model for average preferences and earnings expectations is formulated. This systematic part of the utility, which depends on attributes observable to the researcher and unknown coefficients, is denoted V_{ij}. None of the variables in this systematic part depends on vary between courses within education. Thus, the systematic part varies between applicants and education, but is constant across courses within each education.

Any variation beyond this, e.g. variation between courses within education, variation coming from omitted attributes and from deviations of an applicant's preferences from the average, is modeled as a person course-specific stochastic term, denoted ϵ_{icj}. The two parts are assumed to enter utility additively:

$$U_{icj} = V_{ij} + \epsilon_{icj} \tag{2}$$

The systematic utility is assumed to depend on the log of expected life-time earnings $\log \mathrm{E}LY_{ij}$,[6] earnings risk expressed by the within-education variance of earnings and a nonpecuniary part depending on some vector of variables X_{ij}:

$$V_{ij} = V(\log \mathrm{E}LY_{ij}, \sigma^2_{Y,j}, X_{ij}) \tag{3}$$

Note that expected earnings and nonpecuniary attributes may vary between educations and between individuals within education, while earnings risk is fixed for each education. As noted above, within-education variation enters only through the stochastic term.

The next subsections elaborate on the specification of earnings, nonpecuniary utility and the specification of the random terms, ϵ.

[6] The same functional form is also used by e.g. Beffy *et al* (2011) and Nielsen and Vissing-Jorgensen (2005) studying choice of education. Dagvik *et al* (2006) show that log income has both some theoretical and empirical support as a functional form for the utility of income.

4.2 Expected earnings

The choice of education will depend on the individuals *expected* earnings. In the current setting, expectations are not observed, and must thus be modeled.

Every individual i has some expected earnings $\mathrm{E}Y_{ija}$ in every education j at every age a. These are assumed to be the product of an individual-education-specific constant term, and an education-specific age-earnings profile:

$$\mathrm{E}Y_{ija} = \alpha_{ij}\beta_{ja} \tag{4}$$

Thus, earnings vary between time and between educations, and the individuals have a belief about their relative ability or degree of success in each education. For choice of education, the individuals care about their expected lifetime earnings, which is given as the discounted sum of earnings over the age profile:

$$\mathrm{E}LY_{ij} = \sum_a \delta^a \mathrm{E}Y_{ija} \tag{5}$$

With the multiplicative separability assumed in (4), expected lifetime earnings can similarly be separated into an individual-education-specific factor and an the discounted value of an education-specific earnings-profile:

$$\mathrm{E}LY_{ij} = \sum_a \delta^a \alpha_{ij}\beta_{ja} = \alpha_{ij} \cdot LY_j \tag{6}$$

The first term in (6) is thus the applicant's expectation of own relative earnings potential in a given education, while the second term is the applicant's expected average lifetime earnings for the education.

For the average lifetime earnings, the applicants are assumed to simply use the population averages. As argued by Manski (1993), it is more reasonable that young people are able to observe average earnings than complicated earnings functions. Betts (1996) finds that the single most important source of information on earnings is newspapers and magazines, indicating that the students' knowledge is based on general information.

Thus, the applicants are not assumed to have knowledge of a detailed function determining their relative earnings. Rather, expected relative earnings is assumed to be a simple function of information that can be assumed to be available to the applicants: Some unobserved measure of their absolute ability across all educations and their relative ability and the earnings variance within each education. Relative ability is measured as how many standard deviations the applicant's admission score (G) differ from the average of all student admitted within that education: $\tilde{G}_{ij} = (G_{ij} - \bar{G}_j)/\sigma_{G,j}$ With detailed information on admission requirements, it is reasonable that students have a good idea about their relative academic performance. Furthermore, the variance of earnings may matter for the applicants' expectations. In particular, an applicant of high ability may expect a higher return to that ability in a high-variance education. This is summarised in the following function for earnings expectations:

$$\alpha_{ij} = \exp(\alpha_i + \alpha_1 \tilde{G}_{ij} + \alpha_2 \tilde{G}_{ij} \sigma^2_{Y,j}) \tag{7}$$

Applicants of average absolute ability and with academic performance equal to the average within a given education are assumed to expect earnings equal to the average within that education. Applicants of higher (lower) ability may expect higher (lower) earnings, their expectation increasing with $\alpha_1 + \alpha_2 \sigma^2_{LY,j}$ for each standard deviation increase in admission points. Thus, we should expect $\alpha_1, \alpha_2 \geq 0$.

4.3 Estimation of lifetime earnings and earnings risk

I estimate non-parametric earnings profiles separately for each of the educations, using a ten-year panel data set (1999-2008), allowing for individual-fixed effects:[7]

$$\log Y_{ijt} = \alpha_{ij} + D_{it}\beta_j + \nu_{ijt} \tag{8}$$

[7] The lifetime income calculations are discussed in more detail in Kirkebøen (2010), which also discusses the robustness of the calculations to several assumptions made.

In the earnings equation, Y_{ijt} represents the earning of an individual i, in educational group j at time t. α_{ij} is the individual-fixed effect, D_{it} a vector of indicator variables, representing the experience of i at time t, β_j is a vector of experience effects, which yields the earnings profile, and ν_{ijt} an iid mean zero disturbance term. This earnings equation is consistent with the expectations in (4). Earnings vary flexibly between individuals, and flexibly over time in a way that is shared by all individuals.

Lifetime earnings are calculated for an individual that completes her education at the stipulated age A_j, which is the sum of stipulated duration of the education (S_j) and the school starting age (A). She then starts working and subsequently works every year until retiring when reaching age 67. Thus, at age a she has $a - A_j$ years of work experience. Predicted earnings with a given education of length S_j at a given age a is calculated from the average estimated individual-fixed effect of the group, and the estimated earnings profile:

$$\hat{Y}_j(a) = \begin{cases} \exp\left(\bar{\alpha}_j + \hat{\beta}_{j,a-A-S_j} + \frac{1}{2}(\sigma^2_{\alpha,j} + \sigma^2_{\nu,j})\right) & a > A + S_j \\ Y^0 & a \leq A + S_j \end{cases} \tag{9}$$

Because the log transform is a concave function, by Jensen's inequality, antilog of predicted log earnings underpredicts earnings: $\exp(\mathrm{E}\log Y) \leq \mathrm{E}Y$, with equality only when there is no uncertainty in Y. However, as log earnings is approximately normally distributed I correct for this by adding 1/2 times the residual variance of income, i.e. the sum of the variances of α and ν. For ages at which an individual is not expected to have completed his education, I set earnings to a small, fixed amount, to reflect earnings while studying.

Expected lifetime earnings for an education group can be calculated as the discounted sum of predicted earnings over the life cycle, from graduation from secondary school around age 20 to retirement at 67:

$$\hat{L}Y_j = \sum_{a \in [20,66]} \delta^{(a-20)} \cdot \hat{Y}_j(a) \tag{10}$$

This is the average earnings measure used in (6).

The variance of earnings used in (3) and (6) is the the same as in (9): $\sigma^2_{Y,j} = \sigma^2_{\alpha,j} + \sigma^2_{\nu,j}$. It can be argued that earnings dispersion ($\sigma^2_{\alpha,j}$) and variability ($\sigma^2_{\nu,j}$) have different roles in the determination of expected earnings and choice of education. However, these are highly correlated, and as earnings dispersion is also greater than variability, earnings dispersion is very highly correlated with total variance (coefficient of correlation $> .99$), such that this distinction is of little empirical importance.

As educational choices vary significantly with gender I will do all choice estimations separately by gender. It is not however a priori clear if earnings should be calculated separately by gender. One question is if earnings is reported by gender or as an average across gender in channels the applicants have access to, e.g. media. Another question, particularly relevant for young women, is whether older men or women give the more relevant indication of one's own future earnings, given the changes and convergence between genders in labor force participation over the last decades.

Finally, earnings is estimated from the years 1999-2008. I will proceed to estimate educational choices for application data ranging from 2004 to 2009. Thus, there is a partial overlap between the two data sources, and the applicants in the earliest years can not possibly have known the earnings in the latest years. This is likely to be of little concern, as the lifetime earnings express very persistent differences. Kirkebøen (2010) compare the lifetime earnings calculated from the years 1999-2008 to similar earnings calculated from 1989-1998, finding a coefficient of correlation of about 0.97.

Lifetime earnings is a good measure in a situation with full information and no borrowing constraints, but credit-constrained individuals may care more for initial earnings. Also, the applicants may have higher discount rates than what used in the calculation of lifetime earnings. Berger (1988) finds that a measure of earnings over a longer period explains choice of major better than initial earnings.

Still, as the earnings measure in some sense is arbitrary, the sensitivity to this will be investigated in Section 5.

4.4 Estimation of choice of education

Earlier, the utility from each course is assumed to depend on log expected lifetime earnings, variance of earnings and a vector of nonpecuniary attributes as described by (3), as well as other, nonmodeled variation in the stochastic term.In order to estimate the choice model, I assume that the systematic utility is linear in each of the arguments:

$$V_{ij} = \gamma \log \mathrm{E} LY_{ij} + \eta \sigma_{Y,j}^2 + X_{ij}\theta \tag{11}$$

As for specifying the X_ij vector, there may be several kinds of nonpecuniary attributes, e.g. consumption value of studying and preferences for other career attributes than earnings. As both may vary with field and level of education, I include dummies for field and level in the specifications of utility. The interpretation of the coefficients on these dummies will then capture the average preference for the respective fields and levels, irrespective of whether that utility stems from studying or if the utility is from working after graduation.

Mean and standard deviations of earnings do not fully capture the labor market outcomes associated with an education. To investigate if other attributes influence choices, I control for average time unemployed, average hours of working time per week, and the shares of individuals employed in the public sector and self-employed.

Also, students' choice of field have been shown to vary with parental education. As parental education can only influence choices if it is interacted with attributes of the alternatives, I construct in total four variables that measure similarity in field and squared deviation in duration compared to each of the parents' educations. If applicants want to conform to their parents' educations, we should expect a positive coefficient on similarity in field, and a negative on squared deviation in duration.

Finally, comparative advantage may have a role in explaining educational choices. Paglin and Rufolo (1990) find that the level of quantitative skills is important for education choice and earnings. To investigate this, I interact an indicator variable for whether the education is maths-intensive with indicator variables for whether the applicant has respectively one and two years of elective math in upper secondary.[8] However, as choice sets largely depend on qualifications in maths and science, they are also likely to pick up an element of comparative advantage.

[8]Math-intensive educations are those that mostly consist of courses with formal maths requirements: Business educations, science and engineering, architecture as well as medicine and dentistry, veterinary science and pharmacology. Some of these require two years elective maths, other one year.

However, for choices, only utility differences matter, not utility levels. Therefore, applicants' characteristics cannot themselves influence choices, as all utility comparisons are done between alternatives, within individuals. This means that the person-specific earnings ability from (7) will cancel in utility comparisons. Inserting for the earnings expectations from (6) and (7), this becomes:

$$V_{ij} = \gamma \left(\alpha_i + \alpha_1 \tilde{G}_{ij} + \alpha_2 \tilde{G}_{ij} \sigma^2_{Y,j} + LY_j \right) + \eta \sigma^2_{Y,j} + X_{ij}\theta \tag{12}$$

Thus, comparing two educations j and j' the individual-specific ability (α_i) cancel out: [9]

$$\begin{aligned} V_{ij} - V_{ij'} &= \gamma\alpha_1(\tilde{G}_{ij} - \tilde{G}_{ij'}) + \gamma\alpha_2(\tilde{G}_{ij}\sigma^2_{Y,j} - \tilde{G}_{ij'}\sigma^2_{Y,j'}) \\ &\quad + \gamma(LY_j - LY_{j'}) + \eta(\sigma^2_{Y,j} - \sigma^2_{Y,j'}) + (X_{ij} - X_{ij'})\theta \end{aligned} \tag{13}$$

Choice of education depends on all systematic differences, as well as the stochastic terms, ϵ_{ijc}. An applicant will choose education j if for some $c \in C_{ij}$

$$\epsilon_{ijc} \geq \epsilon_{ij'c'} - (V_{ij} - V_{ij'}) \quad \forall \{c' | c' \in C_{ij'}, j' \in \Omega_i\} \tag{14}$$

An often assumed distribution for the stochastic terms in choice models is iid extreme value, leading to convenient logit choice probabilities. However, there are two reasons to choose a more general distribution in the current application. First, standard logit choice probabilities imply zero correlations between the random terms. This may be unreasonable, as some pairs of courses are very different in content and which careers they qualify for, while other pairs of courses are identical or almost so, for example with the exception of the institution that offer them. Thus, there should be a varying degree of substitutability. Second, the focus of this paper is choice of education among broadly defined alternatives, not the determinants of choice of institution or specific course within education.

The stochastic terms are thus assumed to be independent of the systematic utility, and

[9] The same would of course happen to any characteristic X_i that does not vary between educations. However, if the effect of an characteristic is allowed to vary between the alternatives, i.e. the characteristic is interacted with a alternative-specific constant term in the utility function, the effect will not cancel out from the utility comparisons (except for a normalization, obtained by omitting the characteristic for one attribute).

have a generalised extreme value distribution, i.e. cumulative distribution function

$$\exp\left(-\sum_j \left(\sum_{c\in C_{ij}} \exp(-\epsilon_{ijc}/\rho)\right)^{\rho}\right) \tag{15}$$

This corresponds to a nested logit model (see e.g. Train (2003)). Courses are the lowest-level alternative, while education correspond to nests. The choice of course is decomposed into two choices. Applicants choose education, and course within education. The independence of irrelevant alternatives still holds in the choice of education, and in the choice of course conditional on education, but no longer for the unconditional choice of course. The stochastic term of two courses ϵ_{ijc} and $\epsilon_{ij'c'}$ are uncorrelated if they belong to different educations, i.e. if $j \neq j'$, but if $j = j'$ they are positively correlated, with $1-\rho$ indicating the degree of correlation.[10] If $\rho = 0$ the stochastic terms are perfectly correlated within each education, such that there is no difference between courses within an education. If $\rho = 1$ the stochastic terms are identically and independently distributed across all courses and educations, such that there is no correlation between the stochastic terms within the same education. Thus, there is a fixed correlation within each education.

The probability of interest is that of choosing a given education j, i.e. P_{ij}. Furthermore, because there is no variation in V_{ij} for $c \in C_{ij}$, this becomes an ordinary logit probability, adjusted for the number of courses available in the applicant's choice set, m_{ij} (see Appendix A):

$$\begin{aligned} P_{ij}(\Omega_i) &= \frac{\exp(V_{ij} + \rho \log m_{ij})}{\sum_{j'} \exp(V_{ij'} + \rho \log m_{ij'})}, \\ V_{ij} &= \gamma\alpha_1 \tilde{G}_{ij} + \gamma\alpha_2 \tilde{G}_{ij}\sigma^2_{Y,j} + \gamma L Y_j + \eta\sigma^2_{Y,j} + X_{ij}\theta \end{aligned} \tag{16}$$

In (16) the α_i's are suppressed as these cancel in comparisons, and the dependency of the probability on the applicant's choice set, $\Omega_i = \{j | m_{ij} > 0\}$ is emphasised. As discussed in Section 3 the applicants do not know m_{ij}. They know whether they have the formal

[10] In most presentations, including Train (2003), ρ is allowed to vary between educations. In this paper, it will be constant across all educations. While it could be argued that the degree of correlations in the random terms vary between different education, the relatively large number of educations (19) also make it difficult to estimate separate correlations.

qualifications, but do not know at how many courses they may be admitted. The number m_{ij} depends on their own admission score and those required at the different courses, which in turn depend on the admission scores of the other applicants. The uncertainty in own score is likely to be small. Applicants who are not still in school will know their score, while those still in school are probably able to fairly accurately predict it, based on grades received so far throughout the school year. While the required scores are unknown, last year's required scores are known, and even distributed to the potential applicants, so it seems reasonable that the students calculate m_{ij} based on these. These data are also available for the estimations.

For the estimations m_j is calculated as the number of courses an applicant with the same score could have been admitted to the year before. This approach disregards the difference between the applicants score and last year's requirement, whether the applicant just would (not) have been admitted, and thus could fear (hope for) a small change in the requirement, or if the difference is so large that a change in admission status is unrealistic. However, the m_j's are highly correlated with the average difference between score and requirement. Also, m_j will mostly be from 10-100, such that if the admission requirements are uncorrelated, the large number of specific courses will mean that increases and decreases in admission requirements will cancel out. However, for some educations m_j is much smaller. Also, if m_j varies systematically, e.g. in response to shifts in aggregate preferences for education, there is more scope for a discrepancy between the applicant's expectations and the choice set inferred from the previous year.

The model for educational choice is estimated on unusually rich application data, where each applicant rank up to ten alternatives. Thus, the amount of information is more extensive than in a situation where only the most-preferred choice is known. To fully utilize these data, a rank-ordered logit model is employed. By virtue of the IIA property, excluding any education from the choice set does not alter the ranking of the remaining. Thus, the probability of observing a specific ranking of courses is the probability of having the first choice as the most-preferred from the full choice set, the second choice as the most-preferred in the remaining set excluding the first choice, and so on, i.e. a product of logit probabilities. For an applicant with a choice set of available educations Ω_i, the probability of having the ranking $R_i = \{j, k, l\}$, which means that $j \succ k \succ l \succ$ *all other educations*, is given as:

$$\begin{aligned} P(R_i|\Omega_i) &= P_{ij}(\Omega_i) \cdot P_{ik}(\Omega_{i,-j}) \cdot P_{il}(\Omega_{i,-jk}) \\ &= \frac{\exp(V_{ij} + \rho \log m_{ij})}{\sum_{j'} \exp(V_{ij'} + \rho \log m_{ij'})} \\ &\cdot \frac{\exp(V_{ik} + \rho \log m_{ik})}{\sum_{j' \neq j} \exp(V_{ij'} + \rho \log m_{ij'})} \\ &\cdot \frac{\exp(V_{il} + \rho \log m_{il})}{\sum_{j' \neq j,k} \exp(V_{ij'} + \rho \log m_{ij'})} \end{aligned} \tag{17}$$

Each element $P_{ij}(\Omega_i)$ in (17) is the choice probability in (16), with Ω being the set of educations available to the applicant, i.e. with $m_{ij} > 0$. $\Omega_{i,-j}$ indicates the set of available educations excluding j, i.e. the educations to be considered as a second choice, when the applicant has already ranked j first, and so on. A likelihood function can then be constructed by multiplying the contributions from each individual, given in (17), such that the log likelihood becomes:

$$ll = \sum_i \log \left(P(R_i|\Omega_i)\right) \tag{18}$$

This can be maximized by standard methods to get the MLE of the coefficients in (16).

4.5 Data description

Application data are gathered from SO's centralized registration of applications, for the years 2004-2009. 20 educations are constructed from about 1300 specific courses at different institutions. Table 10 in Appendix B lists the different educations, and how these are classified according to field and level.[11]

Table 1 presents descriptive statistics on the attributes of the educations. Labor market attributes - log earnings, variance of log earnings, unemployment, working time and shares working in the public sector and self-employed - are calculated from administrative register data for the years 1999-2008, that cover the entire working-age population.[12] Whether the

[11] The fields are health and social work, teaching, business and administration, science and engineering, law and social sciences and humanities. Levels are Bachelor, Master and unspecified. While professionally oriented courses have a clear level, broader university studies do not. Students are admitted to a Bachelors course initially, but for most students this is not a final destination, but rather a requirement to enter a Masters course.

[12] There is no common classification of courses or clear link from the application data to other administrative

educations require maths, the number of courses in each education and the share of applicants who have either chosen the relevant education as their most-preferred or ranked it in the application is taken from the application data.

Table 1: Descriptive statistics: Educations

	Mean	SD
Lifetime earnings	12.25	3.111
Earnings dispersion	0.327	0.0628
Earnings variability	0.202	0.0324
Unemployment	0.132	0.0683
Hrs work/week	32.16	1.442
Share in public sector	0.433	0.252
Share self-employed	0.0779	0.107
Requires Math	0.450	0.510
Number of specific courses	54.40	55.76
Share qualifying (at least one course)	0.833	0.303
Share 1st choice	0.0500	0.0414
Share ranked	0.133	0.110
Observations	20	

Although the model in principle is identified as long as the number of education-specific variables is smaller than the number of educations, the small number of educations makes it difficult to empirically separate the effect of many different education-specific variables, as the effect of these is essentially estimated from the average preference for an education. A further challenge is that some of the education-specific variables are correlated, in particular log earnings and the variance of log earnings. This suggests a parsimonious specification of the education-specific share part of the utility, such that not all variables will be used in the main specifications.[13]

A main distinction in the admission to higher education is whether the applicant has elective Math subjects from upper secondary, this - and in some cases further science subjects - is a requirement for several educations, but there are no corresponding requirements for

registers. The analyses presented thus are based on a custom-made link, emphasizing educations that are well-defined in both data sets. These cover 94 percent of the applications and 77 percent of completed higher educations of 30-year olds in 2008.

[13]Furthermore, the total number of variables must be smaller than the total number of observations, i.e., ranked educations. This is of little relevance in this case.

other subjects.[14] Also, Paglin and Rufolo (1990) find that quantitative and verbal ability is a relevant dichotomy, with the former being more highly priced in the labour market. Thus, parsimonious specification for field is whether or not an education requires Math. A more detailed control would be the classification of in fields and level in Table 10 in Appendix B.

Table 2 presents descriptive characteristics of the applicants in Panel A, and of the applicant-education pairs in Panel B. About 37 percent of the applicants are excluded from the analyses because of missing data on admission score, and another 6 percent because of missing data on parental education. The final sample contains data from 301,678 applicant/year observations.[15] From Table 2 we see that about 60 percent of the applicants are female, and that average age at application is 21 years, but with significant dispersion. The main analysis will be restricted to the applicants who are 23 years or younger. One reason is that the estimated lifetime earnings are less relevant as the students become older, as the remaining time in the labour market and thus potential return on investment in education will fall. Also, there is a number of applicants of high age, many already with higher education. It is not clear if these have similar preferences as young applicants making decision for a career.

The average applicant satisfies the formal requirements for almost 16 out of the 19 education. She has a score of 44 points, which means that she can expect to qualify for on average 38 courses within each education. Out of the 10 possible, the average applicant has ranked seven courses, on average within 2.6 different educations.[16]

Looking at the ranked educations in Panel B, in 86 percent of the cases where an applicant satisfies the formal requirements for qualification, she could have been admitted the previous year. For the ranked educations this share is higher, at 95 percent. Thus, it is uncommon, but not unheard of, that an applicant applies for an education she would not be admitted to the previous year.

[14]I.e., if all of an applicants elective subjects from upper secondary are whitin science, she may study science or humanities, while if no elective subjects are not within science she may not study science.

[15]An applicant may have applied in several years.

[16]Obviously, many applicants have ranked several different courses within the same education. However, there is also variation in educations, such that preferences do not seem to be lexicographic, with education dominating.

Table 2: Descriptive statistics: Applicants

Panel A: Individual characteristics

	Mean	SD
Female	0.595	0.491
Age	21.15	2.133
Admission score	44.30	7.529
Educations qualified previous year	16.67	1.753
Courses qualified previous year	35.96	42.16
Number of courses ranked	6.940	2.920
Number of educations ranked	2.651	1.495
Observations	302502	

Panel B: Individual-education match-specific characteristics

	Unranked		Ranked		Total	
	Mean	SD	Mean	SD	Mean	SD
Applicant qualifies	0.844	0.363	0.948	0.223	0.862	0.345
Same field as mother	0.142	0.349	0.128	0.334	0.140	0.347
Same field as father	0.106	0.307	0.108	0.310	0.106	0.308
(Length of schooling - mothers length)^2	18.68	40.06	21.26	42.78	19.12	40.55
(Length of schooling - fathers length)^2	17.45	36.72	19.86	39.24	17.86	37.17
Require Math $\cdot$ applicant has 2 years Math	0.171	0.376	0.198	0.399	0.175	0.380
Require Math $\cdot$ applicant has 3 years Math	0.129	0.335	0.166	0.372	0.135	0.342
Observations	4324456					

Note: In panel B the sample is restricted to educations where the applicant satisfies the formal requirements for qualification.

5 Results

Tables 3 and 4 present results from the estimation of (17) for men and women respectively, with the choice probabilities given from (16). For each applicant, the choice set consists of those educations for which educations the applicant satisfies the formal requirements.

For men, in Table 3, there is a significantly positive impact of log earnings on choice across all specifications. The preferred specification is (4). However, to investigate the sensitivity of the estimated preference to the specification of the control variables, several other specifications are also reported. As it may be difficult to gauge the magnitude of the effects in, I will for the time being focus on the main patterns. I will discuss the magnitudes of the estimated effects in more detail later, in relation to the estimated valuation of different characteristics and the effect on simulated applications.

In specification (4), a strong preference for earnings is found, corresponding to an elasticity (the percent change in number of applicants for a one-percent change in earnings) of about 5. Furthermore, earnings risk is found to be negative, while risk interacted with relative ability is positive. The former is as expected, while the latter is consistent with applicants expecting an earnings premium to relative ability, and that this increases with the variance of earnings, as discussed in Section 4.2.

No available courses, given the applicant's admission score and the previous year's thresholds, is negative, as expected. Also as expected, the number of available courses is positive, with the coefficient indicating a correlation of the unmodeled utility contributions from the specific courses of about $1 - 0.61^2 = 0.63$ within each education. Relative ability enters negatively. This would be surprising if relative ability only captures expected earnings. However, the applicants are likely to also have a direct preference for the ability of their peers, i.e., relative ability also enter among the nonpecuniary variables X_{ij}. Then, a negative coefficient indicates a preference for educations with high ability peer students.

As argued in Section 4.5, whether an education requires Maths is a parsimonious control for field, and this interacted with level of Math from upper secondary may be an indicator of comparative ability. There is a large negative coefficient on Math intensity, indicating a

Table 3: Preferences for expected earnings and nonpecuniary attributes, men

	(1)	(2)	(3)	(4)	(5)	(6)
Log lifetime earnings	2.038***	2.398***	3.402***	4.977***	3.404***	6.076***
	(0.0100)	(0.0164)	(0.0382)	(0.0590)	(0.0510)	(0.0794)
Variance of log earnings			3.501***	-2.200***		-20.50***
			(0.115)	(0.150)		(0.306)
Rel ability · var log earn			2.939***	3.208***		2.916***
			(0.0400)	(0.0414)		(0.0425)
Log available courses		0.613***	0.659***	0.463***	0.536***	0.414***
		(0.00209)	(0.00241)	(0.00373)	(0.00322)	(0.00406)
No available courses		-0.499***	-0.0583***	-0.147***	-0.323***	-0.289***
		(0.0108)	(0.0111)	(0.0121)	(0.0115)	(0.0128)
Relative ability		-0.896***	-1.194***	-0.893***	-0.551***	-0.764***
		(0.00600)	(0.00975)	(0.0114)	(0.00917)	(0.0119)
Requires Math			-1.222***	-1.405***	-1.001***	-1.189***
			(0.0128)	(0.0297)	(0.0289)	(0.0500)
Require Math · applicant has 2 years Math			0.287***	0.174***	0.228***	0.296***
			(0.0161)	(0.0171)	(0.0171)	(0.0170)
Require Math · applicant has 3 years Math			0.497***	0.548***	0.729***	0.521***
			(0.0137)	(0.0138)	(0.0136)	(0.0139)
Unemployment						-3.182***
						(0.214)
Hrs work/week						0.122***
						(0.00739)
Share in public sector						1.227***
						(0.0377)
Share self-employed						8.547***
						(0.108)
Field and level dummies				Yes	Yes	Yes
Parental ed. interactions		Yes	Yes	Yes	Yes	Yes
Log likelihood	-616446.7	-534955.1	-525820.2	-522918.5	-526217.2	-519444.6
Pseudo R^2	0.0318	0.160	0.174	0.179	0.174	0.184
No. of observations	1561368	1561368	1561368	1561368	1561368	1561368
No. of individuals	103951	103951	103951	103951	103951	103951

Note: Estimates of coefficients for the choice model (16), using ranked logit estimation as in (17). Choice sets are all educations for which the applicant satisfies the formal qualification requirements. Field and level dummies control for fields and levels as indicated in Table 10 in Appendix B. Interactions with parental education are two variables reflecting the squared difference in length relative to mother father's education, as well as two variables indicating similarity in field, as shown in Table 2.
Standard errors in parentheses. * $p < 0.10$,** $p < 0.05$, *** $p < 0.01$.

Table 4: Preferences for expected earnings and nonpecuniary attributes, women

	(1)	(2)	(3)	(4)	(5)	(6)
Log lifetime earnings	-0.142***	-0.537***	1.861***	2.698***	-0.434***	3.153***
	(0.00864)	(0.0134)	(0.0312)	(0.0512)	(0.0477)	(0.0638)
Variance of log earnings			-3.320***	-9.796***		-11.92***
			(0.103)	(0.130)		(0.312)
Rel ability · var log earn			3.668***	4.052***		4.119***
			(0.0336)	(0.0356)		(0.0361)
Log available courses		0.394***	0.370***	0.171***	0.304***	0.160***
		(0.00156)	(0.00169)	(0.00282)	(0.00265)	(0.00347)
No available courses		-1.015***	-0.356***	-0.462***	-0.555***	-0.501***
		(0.00911)	(0.00990)	(0.0106)	(0.00963)	(0.0114)
Relative ability		-0.903***	-1.039***	-0.907***	-0.558***	-0.913***
		(0.00520)	(0.00676)	(0.00828)	(0.00767)	(0.01000)
Requires Math			-1.248***	-0.828***	-0.359***	-0.891***
			(0.0107)	(0.0235)	(0.0229)	(0.0393)
Require Math · applicant has 2 years Math			0.429***	0.267***	0.371***	0.277***
			(0.0143)	(0.0154)	(0.0154)	(0.0156)
Require Math · applicant has 3 years Math			0.462***	0.476***	0.726***	0.461***
			(0.0129)	(0.0131)	(0.0129)	(0.0132)
Unemployment						-0.686***
						(0.151)
Hrs work/week						-0.0670***
						(0.00580)
Share in public sector						0.0210
						(0.0318)
Share self-employed						0.376***
						(0.106)
Field and level dummies				Yes	Yes	Yes
Parental ed. interactions		Yes	Yes	Yes	Yes	Yes
Log likelihood	-1043454.6	-969780.4	-954315.5	-925714.9	-934983.4	-925170.5
Pseudo R^2	0.000131	0.0707	0.0855	0.113	0.104	0.113
No. of observations	2170313	2170313	2170313	2170313	2170313	2170313
No. of individuals	156096	156096	156096	156096	156096	156096

Note: Estimates of coefficients for the choice model (16), using ranked logit estimation as in (17). See notes to Table 3.

Standard errors in parentheses. * $p < 0.10$,** $p < 0.05$, *** $p < 0.01$.

strong average dislike. For applicants with two years Math (one elective) there is almost an as strong dislike, while applicants with three years Math are close to indifferent.[17] Note that specification (4) also includes controls for fiel - which is correlated with Math intensity - and level and also for similarity with parents' education. The discussion of these coefficients is postponed.

Specification (1), which gives the estimated preference for earnings, not controlling for any other variables also yields a positive effect, if smaller than in (4). Thus, men on average choose those educations from their choice sets that give higher earnings. Sepeification (2) shows the effect of including several applicant-course-specific variables - number of available courses, availability of any course, relative ability and interactions with parental education. This increases the fit of the model, pseudo R^2 increases from 0.03 to 0.16, but does not change the estimated preference for earnings much relative to (1). Specification (3) shows the results including earnings risk, risk interacted with the applicant's relative ability and a parsimonious specification of field and comparative ability: Whether or not an education requires Math and this indicator interacted with whether the applicant has two or three years Math (one year is compulsory). Adding these controls increase the estimated preference for earnings. This may indicate that applicant would prefer high-earning education, but that to some extent avoid them because high-earning educations also are riskier and often require Math. Surprisingly, the coefficient on risk is positive in this specification.

Specification (5) investigates how excluding risk impacts on the estimated preference for earnings. In (4) there is strong preference for earnings, and aversion for risk. Because these variables are education-specific and correlated, they may be difficult to empirically separate. Compared to (4), the estimated preference for earnings is smaller not controlling for risk in (5), however, there is still a strong preference. Comparing with (3), the preference for earnings is very similar controlling for *either* risk or field and level, and higher when controlling for both (cf. (4)). Finally, specification (6) includes further education-specific covariates: unemployment, average working time and shares in the public sector and self-employed, respectively. This increases the estimated preference for earnings somewhat, and the estimated

[17] Both Math interactions are equal to 1 for applicants with three years Math, such that the sum is the total effect, and Math $\cdot$ 3 years measures the difference relative to applicants with two years Math.

disutility from risk strongly. Unemployment is, unsurprisingly, found to be negative. Note however that earnings already controls for effects of unemployment on earnings, such that this estimate can be interpreted as an effect extends beyond the pure earnings effect. The other three covariates all have positive effects. A preference for long working time is surprising. As for the shares, it is not clear what about the public sector and self-employment that is attractive. Suggestions could be e.g. job security and pension schemes in the public sector, and flexibility for self-employment, but this is speculation. Furthermore, this specification should be interpreted with particular care, due to the problem of empirically separating the effects of a number of different education-specific variables. On this note, it is noteworthy that the coefficients on the education-specific variables (earnings, variance of earnings, Math intensity) generally are more sensitive to the choice of specification than the coefficients on the individual-education-specific variables.

For women, in Table 4, there is a preference for earnings and disutility of risk in the preferred specification (4). However, the preference for earnings is weaker, and the disutility of risk greater, than in the corresponding specification for men. The other coefficients have the same sign as for men, and largely also a similar magnitude. Relative ability interacted with earnings risk has a larger coefficient than for men. Combined with the lower preference for earnings, this implies that the relationship between relative ability, earnings risk and expected earnings (α_2 in Section 4.2) is much higher for women. Furthermore, the number of courses matter less for women, indicating a high correlation within education, about 0.97, and women have a less negative preference for Math.[18] Finally, the model explains less of women's choice of education than for men, as measured by the pseudo R^2.

The estimated preference for earnings depends strongly on control for earnings risk. In the specifications without earnings risk ((1), (2) and (5)), the estimated preference is negative. However, neither the estimated preference for earnings nor for earnings risk change much adding further education-specific controls in specification (6). As for the other education-specific covariates, women show a dislike of unemployment and working time, no significant effect of public sector (conditional on field, which correlates with sector), and a preference for

[18]Note that this is when controlling for field. Not controlling for field, in specification (3), the estimated effect of Math intensity is very similar for both genders.

self-employment, which nevertheless is weaker than for men.

As the nonpecuniary attributes in Tables 3 and 4 do not have the same units, the coefficients are difficult to compare. However, from the estimated coefficients and the assumed utility function (11), it is possible to calculate an compensating earnings change for each of the variables in Tables 3 and 4. To compensate a one-unit change in the nonpecuniary attribute x earnings need to change with $-\theta_x/\gamma$ log points, if θ_x is the coefficient on x.[19] This corresponds to multiplying earnings with $1 + \exp(-\theta_x/\gamma)$. Table 5 reports the relative increase in earnings that would compensate a one-standard deviation change in each of the nonpecuniary attributes reported in Tables 3 and 4. Because of the challenges relating to the education-specific variables Also, it reports the estimated earnings increase associated with the level and field dummies. For the latter, the increase corresponds to a one-unit increase, rather than a one-standard deviation, and compensation is relative to unspecified level and humanities. Thus, with σ_x denoting the standard deviation of x, the compensating earnings in Table 5 is calculated as:

$$CY_x = \begin{cases} \exp(-\sigma_x\theta_x/\gamma) - 1 & \text{for continuous variables} \\ \exp(-\theta_x/\gamma) - 1 & \text{for binary variables} \end{cases} \tag{19}$$

Included in Table 5 is also the standard deviation of log lifetime earnings, corresponding to about 21 percent, for reference. Earnings risk is negative, and to such an extent that across almost all specifications a one-standard deviation increase in risk requires an earnings increase of a 18-20 percent to compensate, corresponding to almost a standard deviation. The exception is the column (1), i.e. specification (4) from Table 3, where there is only a small effect of risk.

As for the similarity with parents' education, similarity in level has a large value for women. A one-standard deviation increase in the squared difference from a parent's education correspond to 10-20 percent earnings decrease, somewhat more for father's education than mother's. The effect is smaller for men (6-7 percent). Also, the effect of same field is about 6-7 percent, for both genders. Women show a strong dislike for Bachelor educations (relative to

[19]This is easily seen by differentiating (11) and setting it equal to zero: $dV = \gamma d(\log \mathrm{E}LY) + \theta dX = 0$.

Table 5: Earning required to compensate for nonpecuniary attributes (share of lifetime earnings)

	Men		Women	
	(1)	(2)	(3)	(4)
Log lifetime earnings	-0.212	-0.212	-0.212	-0.212
	(.)	(.)	(.)	(.)
Variance of log earnings	0.0233***	0.192***	0.208***	0.217***
	(0.00149)	(0.00407)	(0.00426)	(0.00692)
Rel ability · var log earn	-0.163***	-0.124***	-0.339***	-0.302***
	(0.00230)	(0.00214)	(0.00504)	(0.00508)
(Length of schooling - mothers length)^2	0.0639***	0.0629***	0.186***	0.156***
	(0.00275)	(0.00235)	(0.00593)	(0.00509)
(Length of schooling - fathers length)^2	0.0610***	0.0620***	0.176***	0.147***
	(0.00253)	(0.00219)	(0.00554)	(0.00476)
Same field as father	-0.0696***	-0.0566***	-0.0805***	-0.0692***
	(0.00155)	(0.00132)	(0.00277)	(0.00245)
Same field as mother	-0.0421***	-0.0350***	-0.0811***	-0.0699***
	(0.00156)	(0.00131)	(0.00238)	(0.00212)
Bachelor level	0.187***	0.0492***	1.126***	0.925***
	(0.00300)	(0.00372)	(0.0300)	(0.0278)
Master level	0.102***	0.114***	0.264***	0.227***
	(0.00232)	(0.00309)	(0.00595)	(0.00729)
Health and social work	-0.0263***	0.317***	-0.374***	-0.285***
	(0.00278)	(0.0158)	(0.00560)	(0.0121)
Teaching	-0.0313***	0.257***	-0.282***	-0.220***
	(0.00273)	(0.0147)	(0.00513)	(0.0129)
Law and social sciences	0.0360***	0.0808***	-0.0550***	-0.0590***
	(0.00186)	(0.00324)	(0.00383)	(0.00476)
Business and administration	-0.101***	-0.146***	-0.356***	-0.350***
	(0.00474)	(0.00631)	(0.00806)	(0.00926)
Science and engineering	-0.00696	0.0126	0.255***	0.162***
	(0.00486)	(0.00893)	(0.0121)	(0.0171)
Requires Math	0.326***	0.216***	0.359***	0.326***
	(0.00673)	(0.00797)	(0.0109)	(0.0147)
Unemployment		0.0355***		0.0146***
		(0.00271)		(0.00334)
Hrs work/week		-0.0297***		0.0323***
		(0.00200)		(0.00268)
Share in public sector		-0.0486***		-0.00164
		(0.00164)		(0.00249)
Share self-employed		-0.128***		-0.0115***
		(0.00221)		(0.00320)
Observations	1561368	1561368	2170313	2170313

Note: Valuation is calculated as in equation (19), based on the estimates from Tables 3 and 4, specification (4) in columns (1) and (3) and specification (6) in columns (2) and (4). For variables other than same field as parents, educational level and field and requites Math, the valuation presented is that of a one standard deviation change, as per Tables 1 and 2. Educational levels are relative to no unspecified level, fields are relative to humanities.
Standard errors in parentheses. * $p < 0.10$,** $p < 0.05$, *** $p < 0.01$.

unspecified level), men less so. Women also have a stronger dislike of Master educations than men. Women have a clear preference for health and social work, teaching and business and administration (relative to humanities). Men show a different pattern, with a dislike of health and social work and teaching, but a preference for business and administration. However, for men the estimated effects of field are sensitive to the inclusion of further education-specific covariates, this is much less the case for women. Both genders show a dislike of Math. As for the other education-specific variables (unemployment, work time, shares in public sector and self-employed), these are generally found to be of relatively little importance. The single exception is self-employment, for men a one-standard deviation increase in the share self-employed have the same value as a 13 percent earnings increase.

5.1 Magnitude of the effects

As it can be difficult to gauge the size of the effects from the regression coefficients alone, Table 6 presents predicted number of applicants with the different educations as their first choice in different scenarios. The first column presents the observed number of applicants, for reference. In the second column is the predicted number of applicants, based on specification (4) in Tables 3 and 4, relative to the observed numbers in column (1). We see that the predicted figures largely reproduce the main patterns in choice. For most of the educations the ratio of predicted applicants to actual is between 0.7 and 1.3. There are however some notable exceptions. For women, the shares of engineers and civil engineers are strongly underpredicted. For both genders medicine is underpredicted, while maritime education, which is very rare in the data, is grossly overpredicted.

In order to illustrate the significance of earnings and nonpecuniary attributes, columns (3) and (4) presents the predicted number of applicants in two different scenarios: In (3) earnings are equalized across all educations, and in (4) all attributes except earnings are equalized. In both cases applicants are relative to the predicted number underlying column (2). It is immediately clear that both earnings and nonpecuniary attributes play a large role in educational choices. Disregarding earnings give a large reduction for high-earning educations, up to more than 90 percent for mens' applications to medicine and business school, and

Table 6: Simulations

Panel A: Men

Education	(1) # observed	(2) Predicted, ratio	(3) Equal earnings	(4) Only earnings	(5) Teacher +5%	(6) Earnings +5%	(7) Elasticity	(8) Elasticity, incl. risk
Nursing	2070	0.89	2.93	0.37	0.99	1.33	6.69	1.27
Social work	1770	1.04	2.76	0.36	0.99	1.33	6.69	1.15
Physio- and ergotherapy	3775	0.81	1.92	2.10	0.99	1.33	6.58	0.92
Other health	2157	1.12	2.35	0.37	0.99	1.33	6.63	1.28
Kindergarten teacher	1435	1.01	3.84	0.19	0.99	1.33	6.70	1.55
Teachers' college	2321	0.86	2.45	0.29	1.33	1.33	6.55	1.51
Other teaching	4866	1.28	2.20	0.15	0.99	1.31	6.17	0.99
Business school	5463	0.77	0.09	2.32	0.99	1.31	6.13	0.84
Other commerce	11561	0.79	0.41	0.36	0.99	1.30	6.02	1.46
Engineering	4457	0.78	0.37	0.25	1.00	1.29	5.83	1.68
Journalism	2433	1.01	0.46	0.95	0.99	1.33	6.64	0.73
Medicine	2663	0.62	0.06	6.10	1.00	1.29	5.74	0.73
Dentistry et. al.	679	0.95	0.27	6.52	1.00	1.32	6.39	1.41
Civil engineering	9481	0.79	0.17	0.34	1.00	1.24	4.79	0.68
Architecture	2101	0.75	0.56	2.88	0.99	1.34	6.75	0.39
Law	6680	0.84	0.20	8.81	0.99	1.31	6.27	0.84
Science	6751	1.15	0.58	0.20	0.99	1.28	5.61	1.24
Social sciences	19423	1.26	0.70	0.13	0.99	1.23	4.57	0.78
Humanities	13640	1.17	1.59	0.06	0.99	1.27	5.31	0.89
Maritime education	225	2.64	0.77	0.58	1.00	1.34	6.73	2.11

Panel B: Women

Education	(1) # observed	(2) Predicted, ratio	(3) Equal earnings	(4) Only earnings	(5) Teacher +5%	(6) Earnings +5%	(7) Elasticity	(8) Elasticity, incl. risk
Nursing	17171	0.70	1.48	0.64	0.99	1.15	2.99	0.20
Social work	9644	1.08	1.43	0.58	0.99	1.15	3.03	0.02
Physio- and ergotherapy	7687	0.97	1.18	1.31	0.99	1.16	3.13	-0.26
Other health	8264	1.40	1.32	0.64	0.99	1.15	3.02	0.25
Kindergarten teacher	7103	0.87	1.71	0.85	1.00	1.16	3.12	0.69
Teachers' college	5712	0.93	1.36	0.88	1.16	1.16	3.12	0.57
Other teaching	7189	1.47	1.29	0.66	0.99	1.15	3.03	0.02
Business school	2814	0.82	0.26	3.93	1.00	1.16	3.12	0.00
Other commerce	10083	0.71	0.54	2.33	0.99	1.16	3.12	0.84
Engineering	598	0.44	0.55	4.23	1.00	1.16	3.26	1.50
Journalism	3454	1.08	0.57	3.86	0.99	1.16	3.22	-0.48
Medicine	4108	0.55	0.24	1.25	1.00	1.13	2.55	-0.27
Dentistry et. al.	2645	0.72	0.49	1.04	1.00	1.14	2.80	0.32
Civil engineering	2529	0.68	0.37	0.91	1.00	1.14	2.89	-0.00
Architecture	2123	0.98	0.64	6.20	0.99	1.16	3.27	-0.84
Law	9676	0.70	0.38	3.62	0.99	1.16	3.10	-0.17
Science	4020	1.18	0.66	0.72	0.99	1.15	2.96	0.56
Social sciences	30120	1.18	0.72	0.31	0.99	1.12	2.44	-0.01
Humanities	21132	1.14	1.10	0.33	0.99	1.14	2.72	0.02
Maritime education	24	5.38	0.80	6.85	1.00	1.16	3.30	1.86

Note: "# observed" is the number of first choices for each education. "Predicted, ratio" is the ratio of the number of predicted first choice to observed. Predictions are based on specification (4) in Tables 3 and 4. "Equal earnings" is the predicted number of first choices if there were no differences in earnings between the educations relative to the predicted number in column (2). "Only earnings" presents a similar ratio for the for a situation where all differences are eliminated, expect earnings. "Teacher +5%" gives the predicted number of first choices for all educations if the teachers' college got 5% higher earnings, while "Earnings +5%" measures the effect of an exclusive earnings increase of 5% for each of the educations. "Elasticity" gives the earnings elasticity calculated based on the effect of a 5% increase in column (6). Finally, "Elasticity, incl. risk" calculates an elasticity to a simultaneous increase in earnings and risk, where the relative increases in earnings and risk correspond to the cross-sectional relationship between the two.

reductions of more than 80 percent for women's applications to the same educations. Similarly, disregarding differences in other attributes give very large increases for high-earning educations and large reductions for lower-earning educations.

Column (5) presents the effect of a 5 percent increase with teachers' college. This translates to increases of 33 and 16 percent in the number of applicants for men and women respectively, a large increase. This increase is drawn similarly from all other educations, as follows by the IIA property of the conditional logit.[20] Column (6) presents the effect on the number of applicants to each education of an increase of an exclusive 5 percent in the earnings of that education. For men, the calculated increases range from slightly less than 25 percent to above 30 percent, for women from 12 to 16 percent. Based on these increases column (7) shows the calculated elasticities, which are about 6 for men, and about 3 for women. These are arguably large elasticities. One interpretation is that the earnings measure is assumed to reflect the applicants entire working life, and indeed is one that captures persistent differences between educations, such that a 5 percent increase in an education's relative earnings is actually a large change. However, a concern may be if the estimation succeed in separating the effects of earnings and risk. Thus, column (8) presents an elasticity calculated for joint increase in earnings and risk, where the relative increase in the two corresponds to the cross-sectional correlation. We see that for men, this reduces the estimated elasticity, such that it is now about 1. For women the earnings and risk elasticity is essentially zero.

5.2 Stability and robustness

If the estimates are to be used to predict future applications, this requires preferences to be stable over time. Table 7 investigates the stability of the results. Mostly, the estimated coefficients do not vary much between years. Earnings seems to matter somewhat more for men from 2007 onwards, and in 2006-2008 for women. Earnings risk only have a significant negative effect for men in the years 2006-2008, and is more negative for women from 2006 onwards than in 2004-2005. Again, the individual-education-specific variable, in this case relative ability interacted with earnings risk, is more stable than the education-specific ones.

[20] In a situation where one wants to predict the effect of, say, an increase in teachers' earnings, and also care

Table 7: Stability over time

Panel A: Men

	(1) 2004	(2) 2005	(3) 2006	(4) 2007	(5) 2008	(6) 2009
Log lifetime earnings	4.591***	3.875***	4.573***	5.641***	5.287***	5.191***
	(0.141)	(0.148)	(0.148)	(0.146)	(0.141)	(0.153)
Variance of log earnings	0.00980	-0.101	-4.458***	-3.683***	-3.764***	0.179
	(0.367)	(0.369)	(0.358)	(0.371)	(0.374)	(0.401)
Rel ability · var log earn	3.242***	3.067***	2.945***	3.352***	3.484***	2.581***
	(0.107)	(0.104)	(0.104)	(0.106)	(0.0993)	(0.101)
Log likelihood	-84770.5	-90343.5	-86143.0	-86168.5	-89530.1	-84996.6
Pseudo R^2	0.173	0.179	0.184	0.185	0.191	0.168
No. of observations	260548	268786	259816	259481	266301	246436
No. of individuals	17292	17908	17369	17327	17590	16465

Panel B: Women

	(1) 2004	(2) 2005	(3) 2006	(4) 2007	(5) 2008	(6) 2009
Log lifetime earnings	2.157***	2.154***	3.006***	3.356***	3.213***	2.027***
	(0.135)	(0.133)	(0.126)	(0.123)	(0.117)	(0.130)
Variance of log earnings	-7.757***	-7.544***	-11.57***	-10.99***	-13.35***	-10.19***
	(0.333)	(0.323)	(0.310)	(0.308)	(0.320)	(0.374)
Rel ability · var log earn	3.740***	4.082***	3.823***	4.452***	4.366***	3.474***
	(0.101)	(0.0944)	(0.0893)	(0.0863)	(0.0804)	(0.0870)
Log likelihood	-140349.7	-150433.1	-155359.1	-158130.4	-169835.8	-150319.2
Pseudo R^2	0.114	0.117	0.119	0.113	0.113	0.109
No. of observations	340493	354748	365313	369363	389522	350874
No. of individuals	24580	25659	26366	26617	27913	24961

Note: Estimates of coefficients for the choice model (16), using ranked logit estimation as in (17). Sample and control variables as in specification (4), Tables 3 and 4. See notes to Table 3. Standard errors in parentheses. * $p < 0.10$,** $p < 0.05$, *** $p < 0.01$.

As discussed in Section 4 [CHECK!], it is not clear which earnings measure best approximates the applicants expectations and preferences. For example, it is not clear if lifetime earnings or earnings early in the career matter most, if applicants condition their expectations on gender, or if the average earnings is the most relevant statistic. Table 8 presents estimates controlling for different earnings measures. Column (1) reproduces the results from column (4), Tables 3 and 4, while the other columns present results using alternative earnings measures. In column (2) lifetime earnings is calculated by gender. Column (3) presents results using a measure of early-career earnings, predicted earnings at age 30. Columns (4) and (5) present results using measures of respectively "low" and "high" earnings within the relevant education. The earnings measures used in these columns are calculated using the bottom and top 25 percent of the fixed effect distribution within each education. Finally, column (6) presents results using an earnings measure based not on the entire 1999-2008 earnings panel, but rather just the preceding year.

The results are mostly consistent across the different specifications. Also, the model fit is very similar, with pseudo R^2 varying from 0.179-0.180 for men, and from 0.112-0.113 for women. For men, specifications (4), (5) and (6), using predicted earnings at age 50, earnings for the bottom 25 percent and earning for the top 25 percent respectively, have the highest log likelihoods, higher than the baseline specification in (1). For women, only specification (6) have a higher log likelihood. However, for both genders the differences are small, as is visible from the pseudo R^2. Furthermore, the coefficient on log earnings do not differ much between the specifications with highest log likelihoods.

The coefficient that is most sensitive to the choice of dependent variable is that on earnings risk. This is reasonable, as earnings risk may be given a slightly different interpretation in the different models. If applicants do not know where in the education-specific earnings distribution they will end up, the risk is mostly on the upside (downside) when the earnings measure is the bottom (top) earnings. This is consistent with the results that earnings risk is less (more) deterring when the bottom (top) earnings is the earnings measure.

As discussed in Section 4, earnings is based on the years 1999-2008, while the application

about the number of e.g. kindergarten teachers, this is likely to be a severe restriction.

Table 8: Different earnings measures

Panel A: Men

	(1) Baseline	(2) Gender-specific	(3) At 30 yrs	(4) At 50 yrs	(5) Bottom 25pct	(6) Top 25pct	(7) Yearly
Log lifetime earnings	4.977***	6.111***	2.361***	4.012***	5.241***	5.875***	2.359***
	(0.0590)	(0.0705)	(0.0636)	(0.0436)	(0.0574)	(0.0650)	(0.0373)
Variance of log earnings	-2.200***	-7.792***	-0.226	-0.975***	5.611***	-13.88***	0.347**
	(0.150)	(0.185)	(0.177)	(0.141)	(0.132)	(0.231)	(0.144)
Rel ability · var log earn	3.208***	3.051***	2.640***	3.232***	3.247***	3.293***	2.979***
	(0.0414)	(0.0414)	(0.0411)	(0.0412)	(0.0412)	(0.0415)	(0.0414)
Log likelihood	-522918.5	-522612.4	-525874.7	-522283.2	-522349.2	-522331.1	-524557.8
Pseudo R^2	0.179	0.179	0.174	0.180	0.180	0.180	0.176
No. of observations	1561368	1561368	1561368	1561368	1561368	1561368	1561368
No. of individuals	103951	103951	103951	103951	103951	103951	103951

Panel B: Women

	(1) Baseline	(2) Gender-specific	(3) At 30 yrs	(4) At 50 yrs	(5) Bottom 25pct	(6) Top 25pct	(7) Yearly
Log lifetime earnings	2.698***	2.204***	1.857***	2.059***	2.737***	3.235***	1.699***
	(0.0512)	(0.0526)	(0.0496)	(0.0420)	(0.0532)	(0.0568)	(0.0333)
Variance of log earnings	-9.796***	-6.649***	-10.29***	-8.878***	-5.198***	-17.42***	-8.855***
	(0.130)	(0.122)	(0.147)	(0.125)	(0.127)	(0.213)	(0.125)
Rel ability · var log earn	4.052***	3.982***	3.856***	4.020***	4.036***	4.094***	4.011***
	(0.0356)	(0.0359)	(0.0353)	(0.0357)	(0.0356)	(0.0358)	(0.0354)
Log likelihood	-925714.9	-926226.5	-926402.4	-925914.1	-925790.4	-925481.8	-925797.9
Pseudo R^2	0.113	0.112	0.112	0.113	0.113	0.113	0.113
No. of observations	2170313	2170313	2170313	2170313	2170313	2170313	2170313
No. of individuals	156096	156096	156096	156096	156096	156096	156096

Note: Estimates of coefficients for the choice model (16), using ranked logit estimation as in (17). Specification (1) reproduces specification (4), Tables 3 and 4. In specification (2) a gender-specific earnings measure is used. All other earnings measures are pooled for both genders. Specifications (3) and (4) use predicted earnings at age 30 and 50 respectively. Specifications (5) and (6) use predicted earnings from the subsamples with the 25 percent lowest/highest individual fixed effects. Specification (7) use earnings predicted from yearly cross-sections. Sample and control variables as in specification (4), Tables 3 and 4. See notes to Table 3.

Standard errors in parentheses. * $p < 0.10$,** $p < 0.05$, *** $p < 0.01$.

data range from 2004-2009. These periods partly overlap, such that, although the earnings differences are stable over time, applicants early in the period has not had an opportunity to observe the same earnings data. Column (6) presents estimates using lifetime earnings constructed from yearly cross sections of earnings. As is visible from the table, the log likelihood is higher for the baseline results (although, again, the difference is not large). The estimated effect of earnings is smaller than in the baseline specification for both genders. While the difference may reflect that earnings expectations adjust quickly, it is also possible that earnings expectations adjust slowly, and that the different estimates are due to measurement error in expectations of long-term earnings.

Finally, it's again noteworthy that the individual-education-specific variable is more robust to specification.

All results presented so far are restricted to young applicants (19-23 years) and the choice sets studied are determined by their formal qualifications. In Table 9 the sensitivity of the estimates to these choices is investigated.

Specification (1) reproduces the baseline results, specification (4) from Tables 3 and 4, for reference. In specification (2) all educations are included for all applicants, i.e. the choice sets are not restricted to educations for which an applicant is formally qualified. In specification (3) the sample is the same, but formal qualification are controlled for. As expected, being qualified for an education has a large positive effect on the probability of ranking the education. Furthermore, not controlling for qualifications in any way markedly alters the estimated preference for earnings, which is lower for both genders, and even marginally negative for women. This underlines the relevance of explicitly handling the applicants' choice sets. However, the results controlling for qualifications with a dummy variable give similar preferences for earnings, although risk is less deterring. Thus, how information about the choice sets is used seems less important.

Specification (4) restricts the sample more than the in the baseline specification, by removing educations for which an applicant - based on admission score and last year's admission thresholds - can not expect to be admitted. This may be argued to be overly restrictive. While there is no point in applying for an education for which you know you don't qualify,

Table 9: Different estimation samples

Panel A: Men

	(1) Baseline	(2) All courses	(3) All courses	(4) Sufficient points	(5) Ranked courses	(6) P(ranked)	(7) Old
main							
Log lifetime earnings	4.977***	2.866***	4.655***	5.629***	2.045***	5.137***	2.149***
	(0.0590)	(0.0526)	(0.0538)	(0.0775)	(0.0874)	(0.0651)	(0.175)
Variance of log earnings	-2.200***	4.213***	1.604***	-1.697***	-3.505***	-2.179***	3.589***
	(0.150)	(0.135)	(0.136)	(0.166)	(0.235)	(0.172)	(0.429)
Rel ability · var log earn	3.208***	3.359***	3.554***	2.547***	0.457***	3.539***	2.588***
	(0.0414)	(0.0377)	(0.0382)	(0.0456)	(0.0766)	(0.0457)	(0.109)
Formally qualified			1.407***				
			(0.0100)				
Log likelihood	-522918.5	-656816.0	-645845.9	-456894.2	-135560.8	-382420.2	-77028.3
Pseudo R^2	0.179	0.164	0.178	0.181	0.0242	0.232	0.160
No. of observations	1561368	2079020	2079020	1340551	241445	1510248	264967
No. of individuals	103951	103951	103951	103951	100093		18413

Panel B: Women

	(1) Baseline	(2) All courses	(3) All courses	(4) Sufficient points	(5) Ranked courses	(6) P(ranked)	(7) Old
main							
Log lifetime earnings	2.698***	-0.102**	1.798***	0.820***	1.615***	2.667***	0.785***
	(0.0512)	(0.0426)	(0.0442)	(0.0720)	(0.0750)	(0.0589)	(0.175)
Variance of log earnings	-9.796***	-1.716***	-4.064***	-12.39***	-4.333***	-10.49***	-4.525***
	(0.130)	(0.113)	(0.115)	(0.145)	(0.190)	(0.149)	(0.442)
Rel ability · var log earn	4.052***	4.585***	4.581***	4.016***	1.026***	4.341***	3.476***
	(0.0356)	(0.0318)	(0.0322)	(0.0387)	(0.0591)	(0.0391)	(0.105)
Formally qualified			1.308***				
			(0.00957)				
Log likelihood	-925714.9	-1083248.6	-1072627.5	-846697.6	-259371.9	-660374.2	-108514.2
Pseudo R^2	0.113	0.157	0.165	0.0899	0.0155	0.154	0.138
No. of observations	2170313	3121920	3121920	1875816	410961	2141867	327808
No. of individuals	156096	156096	156096	156096	153995		24042

Note: Estimates of coefficients for the choice model (16), using ranked logit estimation as in (17). Specification (1) reproduces specification (6), Tables 3 and 4. Specification (2) includes all educations, also those for which the applicant not satisfies the formal requirements. Specification (3) excludes educations for which the applicant can not expect to be admitted. Specification (4) restricts the sample to those educations actually ranked by the applicant. Specification (5) estimates a conditional logit model for whether an education is among those ranked. Specification (6) restricts the sample to applicants older than 23 years. Unless otherwise indicated, sample and control variables as in See notes to Table 3. Standard errors in parentheses. * $p < 0.10$,** $p < 0.05$, *** $p < 0.01$.

the applicant may hope that the there is less competition when she applies. This is supported by the observation that the coefficient on not expecting admission (from Tables 3 and 4) is much smaller (in absolute value) than the coefficient on formal qualifications in specification (3). Still, restricting the sample has a limited impact on the estimates for men. For women the estimated preference for earnings is reduced, and the risk becomes more important.

The probability giving an education a particular rank may be expressed as the probability of ranking it (any rank) times the probability of a specific rank conditional on ranking. Specifications (5) and (6) separate these. There may be limits to the applicants' knowledge of the earnings with the different educations, or even of an education's mere existence. However, the education actually ranked been actively considered by the applicant, such that the applicants' may be expected to have more information about these. The estimated preference for earnings is lower restricting the estimations to ranked courses. Also, while men still show a stronger preference for earnings, the gender difference is reduced. Also for risk the gender ranking of the estimated preferences is retained, while the difference is reduced. The preferences estimated from the probability of being ranked are very similar to the baseline estimates. That the preferences estimated from the ranked educations and from the probability of being ranked differ goes against the IIA/logit assumption, indicating that this may be problematic.

Finally, specification (7) estimates the preferences of old applicants (older than 23 years). The results indicate that these do indeed differ from the younger ones, both earnings and risk matter less.

5.3 Heterogeneity in preferences

The results so far have been concerned with *average* preferences. However, there may also be important differences between applicants with different characteristics. For example, Saks and Shore (2005) find that wealthier individuals choose riskier careers. Table 11 in Appendix C presents results for each quartile of the applicants' parents' earnings. The coefficients do not differ much with parental earnings. Earnings is found to matter somewhat more for applicants with high-earning parents, the difference is larger for women. Surprisingly, earnings risk is more negative for applicants with high-earning parents, at least for men.

Table 12, also in Appendix C, presents results by quartile of score from upper secondary. There are larger differences along this dimension than what was found for parental earnings, but few clear patterns. For men, earnings seem to matter more for applicants in the middle quartiles, in particular the second. Furthermore, risk is positive, except in the fourth quartile. For women, the preference for earnings is strongest in the highest quartile, as is the estimated effect of risk. The results for women with low score, in column (1) of Panel B, indicate that these have a dislike for earnings and a preferences for earnings risk. This is not credible, and indicates that the model for some reason fails to describe the educational choices of this particular group.

6 Conclusion

How young people's educational choices depend on the earnings prospects of the different educations is a question of both scientific interest and policy relevance. However, nonpecuniary attributes correlated with earnings, and differences in the prospective students' comparative advantages and choice sets make it difficult to estimate this relationship.

This paper finds large effects of earnings, controlling for choice sets and average preferences for some nonpecuniary attributes. Thus, there may be a large scope for influencing prospective students' choice of education by moderate changes in earnings.

Controlling for choice sets, which also are likely to capture comparative advantage, is essential for correct inference. Other attributes are also found to matter strongly, in particular earnings risk. Furthermore, as high-earning educations tend to have larger earnings risk, this will serve to mask the preference for earnings. The strong negative preference for earnings risk also implies a scope for welfare improvements if policies can reduce earnings risk.

The results are mostly robust to earnings measure, while the specification of the choice set have some impact on the results. However, without individual-level variation in earnings and risk, empirically separating the two is challenging. Increasing both earnings and risk reduces or eliminates the estimated effects of earnings on choices.

Men show a much stronger preference for earnings than women, and are less deterred by risk. There are few clear difference by the applicants' parental income or academic performance.

References

Arcidiacono, P. (2004). Ability sorting and the returns to college major *Journal of Econometrics* Vol 121

Arcidiacono, P., V. J. Hotz and S. Kang (2011). Modelling college major choices using elicited measures of expectations and counterfactuals *Journal of Applied Econometrics* (forthcoming)

Beffy, M., D. Fougére and A. Maurel (2011) Choosing the Field of Study in Post-Secondary Education: Do Expected Earnings Matter? *Review of Economics and Statistics* (forthcoming)

Berger, M. C. (1988). Predicted future earnings and choice of college major *Industrial and Labour Relations Review* Vol 41 No 3

Betts, J. R. (1996) What Do Students Know about Wages? Evidence from a Survey of Undergraduates *Journal of Human Resources* Vol 31

Boskin, M. J. (1974). A Conditional Logit Model of Occupational Choice *The Journal of Political Economy* Vol 82 No 2 Part 1

Boudarbat, B. (2008) Field of study choice by community college students in Canada *Economics of Education Review* 27

Boudarbat, B. and C. Montmarquette (2007) Choice of fields of study of Canadian university graduates: The role of gender and their parents' education IZA DP no. 2552

Dagsvik, J. K., S. Strøm and Z. Jia (2006) Utility of income as a random function: Behavioral charatcterization and empirical evidence *Mathematical Social Sciences* Vol 51

Desposato, S. (2005) A Model of Unorder Multiple Choice with Unobserved Choice Set Selection UCSD Working paper

Dominitz, J. and C. Manski (1996) Eliciting student expectations of the returns to schooling *Journal of Human Resources* Vol 31

Flyer, F. A. (1997) The influence of higher moments of earnings distributions on career decisions *Journal of Labor Economics* Vol 15 No 4

Kirkebøen, L. J. (2010) Forskjeller i livsløpsinntekt mellom utdanningsgrupper Report 2010/43, Statistics Norway

Manski, C. (1993). Adolescent econometricans: How do youth infer the returns to schooling? In *Studies of Supply and Demand i Higher Education* Clothfelder C. and M. Rothschild (eds) Univeristy of Chicago press: Chicago, IL

Manski, C. and D. Wise (1983) College choice in America Harvard University Press: Cambridge, MA

Montmarquette, C., K. Cannings and S. Mahseredjian (2002) How do young people choose college majors? *Economics of Education Review* 21

Nielsen, H. S. and A. Vissing-Jorgensen (2005) The impact of labor income risk on educational choices: Estimates and impied risk aversion Mimeo, University of Aarhus, Denmark

Paglin, M. and A. M. Rufolo (1990). Heterogeneous Human Capital, Occupational Choice and Male-Female Earings Differences *Journal of Labor Economics* Vol 8 No 1

Rochat, D. and J.-L. Demeulemeester (2001). Rational chioce under unequal constraints: the example of Belgian higher education *Economics of Education Review* 20

Saks, R. E. and S. H. Shore (2005) Risk and Career Choice *Advances in Economic Analysis and Policy* Vol 5, Iss 1

Train, K (2003). Discrete Choice Methods with Simulation Cambridge Univeristy Press: Cambridge, United Kingdom

Willis, R. J. and S. Rosen (1979). Education and Self-Selection *The Journal of Political Ecnonomy* Vol 87 No 5

Wiswall, M. and B. Zafar (2011). Determinants of College Major Choice: Identification Using an Information Experience Staff Report no. 400, Federal Reserve Bank of New York

Zafar, B. (2009). College Major Choice and the Gender Gap Staff Report no. 364, Federal Reserve Bank of New York

Zafar, B. (2011). How do College Students Form Expectations? *Journal of Labor Economics* (forthcoming)

A Going from nested logit to (16)

Following e.g. Train (2003), in a nested logit model, the probability of an applicant i choosing course c within education j is can be expressed as $P_{ijc} = P_{ic|j} \cdot P_{ij}$. The systematic utility of course c in education j is V_{ij}, i.e., independent of c. To see the equivalence to standard expositions of nested logit models we can add an element $V_{ic|j} = 0$. This is the contribution from the course to the systematic utility, conditional on the education, such that the systematic utility becomes $V_{ij} + V_{ic|j}$. With $\delta_{ic} = 1$ if a course c is available to an applicant, and zero otherwise, we can then express the probability P_{ijc} as:

$$P_{ijc} = \frac{\delta_{ic} \exp(V_{ic|j})}{\sum_{c' \in C_{ij}} \exp(V_{ic'|j})} \cdot \frac{\exp(V_{ij} + \rho I_{ij})}{\sum_{j'} \exp(V_{ij'} + \rho I_{ij'})} \tag{20}$$

The first fraction is $P_{ic|j}$, the probability of choosing course c from C_{ij}, i.e. conditional on having chosen education j. This probability depends on attributes of the courses, that enter the utility $V_{ic|j}$. The second term is the probability of interest, that of choosing education j. This depends on attributes of the educations that do not vary between courses within each education, entering the term V_{ij}. However, the choice of an education will also depend on the courses that constitutes it. This is captured by the term I_{ij}, the inclusive value of education j, which is a function of the utility of the available courses within the education:

$$I_{ij} = \log \sum_{c \in C_{ij}} \exp(V_{ic|j}/\rho) = \log m_{ij} \tag{21}$$

The last equality follows because all $V_{ic|j} = 0$, and $m_{ij} = \sum_{c \in C_{ij}} 1 = \sum_{c \in C_j} \delta_{ic}$.

B Classification of educations

Table 10: Classification of educations

Bachelor level
Health and social work: Nursing; Social work; Physiotherapy and ergotherapy; Other health
Teaching: Kindergarten teacher; Teachers' college; Other teaching
Business and administration: Business school; Other commerce
Science and engineering: Engineering; Maritime education
Law and social sciences: Journalism
Master level
Health and social work: Medicine; Dentistry, Veterinary Science and Pharmacology
Science and engineering: Civil Engineering; Architecture
Law and social sciences: Law
Unspecified level
Science and engineering: Science
Law and social sciences: Social sciences
Humanities: Humanities

C Result tables

Table 11: Estimated preferences by quartile of parental earnings

Panel A: Men

	(1) 1st quartile	(2) 2nd quartile	(3) 3rd quartile	(4) 4th quartile
Log lifetime earnings	4.855***	4.699***	4.784***	5.510***
	(0.168)	(0.160)	(0.148)	(0.123)
Variance of log earnings	-1.492***	-1.426***	-2.683***	-4.721***
	(0.408)	(0.388)	(0.368)	(0.340)
Rel ability · var log earn	3.159***	3.007***	2.814***	3.050***
	(0.110)	(0.107)	(0.101)	(0.0917)
Log likelihood	-74292.6	-82102.3	-89112.9	-100747.0
Pseudo R^2	0.163	0.167	0.182	0.213
No. of observations	216348	236805	264271	314610
No. of individuals	14909	16131	17648	20063

Panel B: Women

	(1) 1st quartile	(2) 2nd quartile	(3) 3rd quartile	(4) 4th quartile
Log lifetime earnings	2.615***	2.220***	2.362***	3.293***
	(0.137)	(0.133)	(0.123)	(0.109)
Variance of log earnings	-10.07***	-9.757***	-10.16***	-11.82***
	(0.353)	(0.338)	(0.319)	(0.303)
Rel ability · var log earn	3.824***	4.093***	4.086***	3.823***
	(0.0854)	(0.0863)	(0.0865)	(0.0829)
Log likelihood	-156150.5	-168710.5	-159673.7	-147576.1
Pseudo R^2	0.105	0.105	0.116	0.138
No. of observations	361190	382286	370798	360798
No. of individuals	26558	27907	26497	24895

Note: Estimates of coefficients for the choice model (16), using ranked logit estimation as in (17). Sample and control variables as in specification (6), Tables 3 and 4. See notes to Table 3.
Standard errors in parentheses. * p < 0.10,** p < 0.05, *** p < 0.01.

Table 12: Estimated preferences by quartiles of score

Panel A: Men

	(1) 1st quartile	(2) 2nd quartile	(3) 3rd quartile	(4) 4th quartile
Log lifetime earnings	4.824***	7.335***	5.663***	4.215***
	(0.238)	(0.232)	(0.193)	(0.0955)
Variance of log earnings	8.537***	8.726***	2.802***	-6.700***
	(0.474)	(0.398)	(0.379)	(0.344)
Rel ability · var log earn	5.015***	9.463***	7.903***	3.086***
	(0.179)	(0.254)	(0.274)	(0.145)
Log likelihood	-119442.0	-124181.1	-127409.5	-144350.9
Pseudo R^2	0.166	0.179	0.184	0.225
No. of observations	363935	372712	374573	450148
No. of individuals	27972	25905	24153	25921

Panel B: Women

	(1) 1st quartile	(2) 2nd quartile	(3) 3rd quartile	(4) 4th quartile
Log lifetime earnings	-1.224***	1.601***	1.387***	2.058***
	(0.257)	(0.202)	(0.161)	(0.0788)
Variance of log earnings	-2.680***	-1.243***	-1.702***	-8.472***
	(0.634)	(0.405)	(0.360)	(0.276)
Rel ability · var log earn	4.716***	5.006***	4.152***	2.982***
	(0.176)	(0.205)	(0.200)	(0.110)
Log likelihood	-178486.1	-219131.9	-248754.1	-270032.1
Pseudo R^2	0.116	0.106	0.113	0.146
No. of observations	429365	515463	579521	645964
No. of individuals	34506	38826	41547	41217

Note: Estimates of coefficients for the choice model (16), using ranked logit estimation as in (17). Sample and control variables as in specification (6), Tables 3 and 4. See notes to Table 3.
Standard errors in parentheses. * $p < 0.10$,** $p < 0.05$, *** $p < 0.01$.