DISCUSSION PAPER SERIES

IZA DP No. 1979

Social Class and Undergraduate Degree Subject in the UK

Massimiliano Bratti

February 2006

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Social Class and Undergraduate Degree Subject in the UK

Massimiliano Bratti
University of Milan, CHILD
and IZA Bonn

Discussion Paper No. 1979
February 2006

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
Email: iza@iza.org

IZA Discussion Paper No. 1979
February 2006

ABSTRACT

Social Class and Undergraduate Degree Subject in the UK*

Although past research has found strong social class effects on the decision to undertake higher education in the UK, there is only sparse empirical work investigating social class influences on the choice of degree subject at the undergraduate level. Using Universities' Statistical Record data for the period 1981-1991, we find no social class effect on students' undergraduate degree subjects enrolled. Our analysis suggests that in a period pre-dating the mass expansion of higher education, the replacement of student grants with student loans and the introduction of undergraduate student tuition fees, the UK university system granted equal opportunities to students from different social classes in terms of the degree subject enrolled.

JEL Classification: C25, I21

Keywords: degree subject, social class, UK, undergraduate

Corresponding author:

Massimiliano Bratti
DEAS
Università degli Studi di Milano
Via Conservatorio 7
20122 Milano
Italy
Email: massimiliano.bratti@unimi.it

* I acknowledge both the Universities' Statistical Record (USR), as the original depositor, and the UK Data Archive for the use of the data set SN: 3456 Universities' Statistical Record, and financial assistance from the ESRC. This paper is the second chapter of my PhD Thesis (University ofWarwick) and was partly written when I was visiting the Center for Economic Research (CentER) at Tilburg University as a Marie Curie Fellow within the ENTER exchange program. Funding from the European Commission is gratefully acknowledged. This paper benefited from presentations at the University of Warwick, the 2003 IZA Summer School in Labor Economics (Amersee), the II Mediterranean Summer School in Theoretical and Empirical Economics 2003 (Palma de Mallorca), the European Association of Labour Economists 2004 Conference (Lisbon) and the Third Labour Economics Workshop "Brucchi Luchino" (Florence, 2005). I am particularly indebted to Mark Stewart, Jeremy Smith, Derek Leslie, Martin Andrews, Jan van Ours, Qin Tu, Marton Csillag and Ian Walker for useful comments on early drafts of this paper. The usual disclaimer applies.

Introduction

Widening access to education is commonly viewed as an effective way of promoting higher intergenerational mobility and the study of educational attainment has received growing attention by economists. This is especially true for the UK where the empirical analysis of educational attainment using micro data has a long tradition. Several studies have investigated the determinants of the level of education achieved (the number of years of schooling or the highest educational qualification) using cross-section or longitudinal micro data. Some examples include Rice (1987), Micklewright (1989), Ermisch and Francesconi (2001) and Chevalier and Lanot (2002) among others. While previous studies have shown a substantial amount of correlation between parents' and children's incomes and education,[1] there is much less empirical evidence on the effect of family and social background on the choice of subject at tertiary level. However, in a period of increasing access to education, a great deal of the variation in individuals' labour market outcomes (employment opportunities and earnings) may be determined by the *type* in addition to the *level* of education achieved. Empirical evidence that supports the importance of field of study as one of the main determinants of graduates' performance in the labour market is provided by several studies. Smith *et al.* (2000) and Bratti *et al.* (2004), for instance, report significant differences in first destinations of graduates from different subject fields. Large differences also exist in graduates' earnings by degree subject, as shown by Blackaby *et al.* (1999), Walker and Zhu (2001) and Chevalier *et al.* (2002) among others.

Moreover, there is some concern about the lack of workers in high demand fields, such as graduates in computer sciences and IT (see Mason, 1999). Hence, a deeper understanding of the mechanism driving students' choices is important also to explain some apparent inefficiencies in the labour market and to forecast future labour market trends.

Despite the potential interest of the topic, to the best of our knowledge, to date there exists only one empirical study on undergraduate field of graduation in the UK, van de Werfhorst et al. (2003), which analyses survey data for the 1958 British cohort. In the current paper, we aim to contribute to the existing literature

on subject choice in the UK by extending the analysis to several cohorts of university students (from 1981 until 1991), using administrative individual-level data. Unlike many previous studies, we use cohorts of entrant students rather than cohorts of students leaving with a university qualification and model subject choice allowing for a non-zero correlation across the unobserved factors which might simultaneously affect the utilities received from studying different disciplines. Furthermore, using several cohorts of university students we also analyse the changes in social class effects over time.

The paper is organised as follows. Section 1 gives an outline of the reasons why social class should matter for the choice of degree subject. Section 2 briefly surveys the existing literature. Section 3 describes the econometric model and section 4 discusses the main features of the data set and the sample used in our estimates. Section 5 reports the empirical results and Section 6 concludes.

1 Social class influences on the choice of field of study

The case for strong social class influences on educational choices has been put forward by both sociologists and economists. Although the theoretical work has almost exclusively focused on the level of education, the existing analytical framework can also be applied to the choice of field of study.

Starting from the sociological literature, Boudon's (1974) model of 'rational action' states that educational choices depend on the perception of the costs and benefits of each educational alternative available. For the choice of subject field, a related hypothesis elaborated by Kelsall *et al.* (1972) is that low social class students may be more inclined to choose subjects that offer better labour market prospects. This could happen because future labour market outcomes depend more on subject studied for low social class than for high social class students. The latter are likely to enjoy good labour market outcomes once they get a university degree irrespective of their field of graduation, thanks to 'family networks'. Moreover, Kelsall *et al.* (1972) also maintain that low social class students may tend to choose

technical fields of study, which are closer to the occupational experience of many manual working class parents. Boudon's (1974) model closely resembles the one commonly used by economists: Becker's (1975) human capital theory. According to Becker (1975) the costs and the returns of education are the main factors driving educational choices. Some fields might be more closely linked to professions for which the presence of 'social networks' (to which high social class students are typically better connected) is more important to ensure labour market success and a higher economic return of the educational investment. Then, the expected return from different educational fields may differ accordingly over social classes. Previous research has shown the existence of family networks effects. Hansen (2001), for instance, found that the impact of social class on the economic rewards of education varies across educational fields and tends to be largest in 'soft' educational fields (such as social studies and humanities). Moreover, since in the presence of capital market imperfections low social class individuals might have higher costs of enrolling in HE,[2] standard economic theory predicts that these individuals will require a higher return from their investment in university education. The higher return can be obtained by enrolling in subjects highly rewarded in the labour market. Therefore, low social class students will choose relatively 'high performing' subjects (in terms of earnings or employment prospects), a prediction very similar to that elaborated by Kelsall *et al.* (1972). On the grounds of these considerations we advance the following hypotheses:

Hypothesis 1. Low social class individuals are relatively more inclined to enroll in 'technical' degree subjects than middle and high social class students (Kelsall *et al.* 1972);

Hypothesis 2. Low social class individuals are more likely to enroll in subjects which offer better labour market prospects, i.e. higher wages or better employment opportunities, than middle and high social class students (Kelsall *et al.* 1972 and Becker 1975).

Bourdieu's (1984) 'cultural reproduction hypothesis' emphasizes the role of education as a means of reproducing social class. In this context education is the instrument through which the high and middle social classes prevent individuals from lower social backgrounds from accessing the highest positions in society. Two

very important concepts for educational choices are those of 'economic capital' and 'cultural capital'. As stressed by van de Werfhorst *et al.* (2001), according to this hypothesis people from the economic elites prefer lucrative fields, which can ensure a comfortable life, while people from the cultural elites are less interested in economic returns to education and prefer fields in which they can acquire 'cultural capital'. Then the following hypothesis can be put forward:

Hypothesis 3. High social class students are relatively more likely to enroll in subjects which provide cultural capital, such as arts and humanities (Bordieu 1984).

This last hypothesis is also in line with hypothesis 2, since arts and humanities are usually not very highly remunerated subjects.

We may expect that with an increased access to higher education the advantages of having a degree for the upper class are progressively lost, and the type of degree possessed becomes an important distinguishing factor for labour market and social success. Highly prestigious professions, such as the legal and medical ones, offer socially and economically advantageous positions in society, and control over these professions becomes a valuable asset for the upper class. A theory leading to similar conclusions exists also in the domain of economics: the 'social networks' model elaborated by Montgomery (1991). Montgomery (1991) builds a theoretical model to explain the large use of employee referrals as a device for screening job applicants, starting from the observation that workers tend to refer others who are similar to themselves (see Doeringer and Piore, 1971). His model explains "why workers who are well connected (possessing social ties to those in high-paying jobs) might fare better than those who are poorly connected and why firms hiring through referral might earn higher profits" (Montgomery 1991, p. 1414). We may expect that individuals from high social classes are better connected to people working in high-paying jobs, their parents *in primis*. Then we derive the following hypothesis:

Hypothesis 4. High social class students are relatively more likely to enroll in prestigious subjects, i.e. subjects leading to highly paid and often entry-regulated professions, such as medicine and law. In the latter, the comparative advantage of high social class individuals may stem from a direct control over the entry of related professions (Bordieu 1984), or from the existence of social networks effects

(Montgomery 1991).[3]

It is worth noting that the last hypothesis is in apparent contrast with hypothesis 2 since graduates in medicine or law are usually very well paid. However, in reality students may enroll in the subjects which ensure them the highest expected returns conditional on their social class origins.

In the empirical analysis, we shall compare the empirical evidence with the hypotheses outlined above and see which ones seem to better fit the data. Although the hypotheses above suggest why social class may be important for the choice of degree subject, they offer at the same time some counter arguments on the reasons why we could find that social class differences were not significant in the period under study. Some of these reasons could be:

1. individual preferences (or non-pecuniary costs) are not shaped by social class but by factors unrelated to it. One such factor might be the performance in specific subjects at secondary school;

2. given the increasing demand for graduates in the period under study, labour market outcomes did not depend on social class, i.e. family networks were not important and a university degree was sufficient to ensure a good labour market outcome to all graduates (absence of heterogeneous returns to degree subjects according to social class);

3. individuals had in the period under study the same (pecuniary) costs of enrolling in different subjects. This might be the case since there were no tuition fees for home students and student financial assistance was based on means-tested maintenance grants for low-income students.[4]

Thus all factors above may have contributed to making very similar the behaviour of individuals with different social origins.

2 Previous empirical research

In this section, we report a brief survey of the empirical research that investigated university students' choice of field of study at the undergraduate level and which has also analysed the role of social class influences. To the best of our knowledge,

there are only a handful of studies which have investigated social class effects using individual-level data.

Oosterbeek and Webbink (1997) used data from the Netherlands in order to analyse the decision whether or not to attend technical studies. The authors found that children from high income families were less likely to enroll in technical fields, but more likely to persist in their choice once they had undertaken a technical education.

Davies and Guppy (1997) analysed the choice of field of study using US micro data. They found that males were more likely than females to enroll in lucrative fields of study. Moreover, high ability individuals and low social class individuals were more likely to enter high-return fields.

Van de Werfhorst *et al.* (2001) using Dutch data found that children of the cultural elite tended to choose fields where they could acquire 'cultural capital', i.e. non technical fields, while students from the economic elite were under-represented in cultural fields (such as arts and humanities). By contrast, low social class individuals were over-represented in economics and engineering, i.e. lucrative fields.

Rochat and Demeulemeester (2001) investigated the process of study field choice using Belgian data. They found that students with fathers in 'elite' occupations, such as managers, civil servants or professionals, were relatively more likely to enroll in short cycle artistic and pedagogical studies and long cycle curricula in engineering and less likely to enroll in long cycle business, economics and social studies.

Montmarquette *et al.* (2002) estimated a multinomial logit model of subject choice using Canadian micro data. They found no effect of having a parent in a professional occupation, but that students supported by an educational loan were more likely to choose those fields (education or liberal arts) in which the probability of success was higher on average.

We are aware of only one paper investigating field of study at the undergraduate level in the UK with an emphasis on social class and that is by van de Werfhorst *et al.* (2003). The authors analysed the educational choices and the educational performance of the 1958 British Cohort using data from the National Child Development Study. Since there is only this study for the UK, it is worthwhile to devote some space to summarizing the main findings. The focus of van de Werfhorst *et*

al. (2003) is on the role of social class, cultural and economic capital and ability on the subject choice in secondary and tertiary education in Britain. We comment here only on the results relating to university education. Van de Werfhorst *et al.* (2003) estimated a multinomial logit model of subject of graduation considering six broad subject categories and including among the explanatory variables family background variables (such as parental social class and measures of 'economic capital' and 'cultural capital'), ability (verbal and mathematical ability), and measures of comparative advantage (based on O-level subjects choice and performance). The authors found that children from professional backgrounds preferred faculties of medicine and law, even after controlling for ability at age 11 and exam performance at age 16. However, they did not find other social class differences, which, as they pointed out, is not due to the controls for various sorts of school attainments since a model without age-11 and age-16 attainments also shows no other social class effects. However, the authors themselves stated that the lack of a strong social class effect might be due to the specific characteristics of the cohort studied. In fact, at the time of the study only a very small minority of the working class entered HE, and this could be considered as a very particular and selected group (e.g., in terms of academic ability).

3 The econometric model

We assume that an individual can choose a subject group j among three different alternatives, $j \in \{0, 1, 2\}$, which will be defined in section 4.2, each of them providing a utility of:

$$U_{0i} = \beta_0' X_i + \epsilon_{0i} \quad (1)$$

$$U_{1i} = \beta_1' X_i + \epsilon_{1i} \quad (2)$$

$$U_{2i} = \beta_2' X_i + \epsilon_{2i} \quad (3)$$

where i is the subscript for individuals. X_i is the vector of all individual observed characteristics affecting the utility of each group and ϵ_{0i}, ϵ_{1i}, ϵ_{2i} the unobserved components (errors) in these utilities. A possible way of modeling the

choice is to use a multinomial logit model (MNL). However, a strong assumption of the MNL is the independence of irrelevant alternatives, i.e. that the error terms of the utilities associated with the alternatives are uncorrelated, an assumption which contrasts with what we have seen in the previous section. However, some factors affecting subject related utilities are often unobservable, and enter the error terms ϵ_{ji}'s thus generating a correlation between them. This typically happens when using administrative data, which do not contain important information on family background information, which is instead collected through survey data. Therefore, in our case, it is evidently advantageous to use a trinomial probit model (TNP) since the covariance matrix of the error terms is unrestricted. As suggested by Bunch (1991) and Dansie (1985), among others, a convenient way of achieving identification of the TNP is by normalizing one of the utilities to zero. This reduces the dimensionality of the problem. By normalising to zero the utility of the first alternative (U_{0i}), our model becomes:

$$U_{0i} = 0 \tag{4}$$

$$U_{1i} = \beta_1' X_i + \epsilon_{1i} \tag{5}$$

$$U_{2i} = \beta_2' X_i + \epsilon_{2i} \tag{6}$$

where:

$$\begin{pmatrix} \epsilon_{1i} \\ \epsilon_{2i} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{pmatrix} \right).$$

This model represents a formally identified TNP.[5] As observed by Heckman and Sedlacek (1985) the TNP is identified so long as X_i contains a single regressor that varies across individuals and no exclusion restrictions are required for formal identification. However, as stated by Keane (1992) the TNP may suffer from 'tenuous' identification, and exclusion restrictions may contribute to improving the model identification. The problem is likely to arise especially when considering the choice among a number of alternatives higher than three. However, as often happens in labour economics applications, our data set does not contain alternative-specific variables, so no natural exclusion restrictions are available. Since theory does not suggest any obvious exclusion restrictions for non-alternative-specific variables, i.e.

variables affecting the utility of a specific alternative only, we can only estimate the formally identified TNP without exclusion restrictions.[6]

The probabilities of the different outcomes are:

$$\begin{aligned} P(Y=0) &= P(U_0 > U_1, U_0 > U_2) \\ &= P(\epsilon_1 < -\beta_1' X, \epsilon_2 < -\beta_2' X) \\ &= \Phi_2(-\beta_1' X, -\beta_2' X | \rho_{12}) \end{aligned} \tag{7}$$

$$\begin{aligned} P(Y=1) &= P(U_1 > U_0, U_1 > U_2) \\ &= P(\epsilon_1 < \beta_1' X, \epsilon_1 - \epsilon_2 < -\beta_1' X - \beta_2' X) \\ &= \Phi_2\left(\beta_1' X, \frac{\beta_1' X - \beta_2' X}{\sqrt{2-2\rho_{12}}} | \frac{1-\rho_{12}}{\sqrt{2-2\rho_{12}}}\right) \end{aligned} \tag{8}$$

$$\begin{aligned} P(Y=2) &= P(U_2 > U_0, U_2 > U_1) \\ &= P(\epsilon_2 < \beta_2' X, \epsilon_2 - \epsilon_1 < -\beta_2' X - \beta_1' X) \\ &= \Phi_2\left(\beta_2' X, \frac{\beta_2' X - \beta_1' X}{\sqrt{2-2\rho_{12}}} | \frac{1-\rho_{12}}{\sqrt{2-2\rho_{12}}}\right) \end{aligned} \tag{9}$$

where we have omitted the subscript for the individual and $\Phi_2(x_1, x_2|\rho_{12})$ is the bivariate standard normal distribution of the two normal random variables ϵ_1 and ϵ_2 computed at the values x_1 and x_2, respectively, with correlation coefficient ρ_{12}.

4 Empirical analysis

In the present section we describe the data set used, the choice of explanatory variables and the econometric model.

4.1 Data set

In this paper we use individual-level data from the Universities' Statistical Record (USR). The USR was the institution in charge of the collection of the statistical

returns from all university institutions in Great Britain which formerly received Exchequer grants from the University Funding Council (UFC), together with corresponding institutions for the Queen's University of Belfast and the University of Ulster. The USR has stored data from the academic year 1972/1973 until 1993/1994 when it was replaced by the Higher Education Statistics Agency (HESA). HESA data are generally comparable to USR data. However, the new data do not include performance by type of A-level, which is likely to be a very important predictor of students' selection into fields of study, but provides only A-level score in the best three passes. At the time of the study we had no access to HESA data, and despite being a bit old, USR data provide nonetheless the picture of the UK university system during a period pre-dating the big expansion of university education and the steady shift in the burden of funding higher education away from the taxpayer and towards students and their families that took place in recent years. Hence, the analysis in this paper is limited to USR data and 'old' universities, i.e. to the institutions with a university status before the abolition of the binary divide between universities and polytechnics that took place in 1992. The USR data are rich in information concerning the academic life and prior educational qualifications of students and include entire cohorts of students leaving Universities each year.

Using different cohorts of university leavers, it is possible to re-construct the cohorts of entrant students in each academic year. We did it for the academic years 1981-1991. Hence, unlike other papers which use samples of university leavers, in particular students leaving with a university qualification, in this paper we use cohorts of entrant students. We think that this is more appropriate to investigate factors related to the choice of degree subject. Indeed, obtaining a degree in a specific subject is only the final outcome of several processes, namely the choice to enroll in a certain field, that of remaining in the same field along the course, and that of students' progression. Therefore, it is difficult to isolate the effect of the explanatory variables on each of these single processes by analysing cohorts of leaving students. In this paper, we aim to analyse the first subject in which students enrolled and accordingly use cohorts of entrant students.

We are aware of the fact that observing an individual enrolled in a certain field implies that he/she has received an offer by a university, and therefore that also the

supply side is important. However, we have no individual-level data on the subject preferences stated by students at the application stage. We observe only students' revealed preferences. In this sense, we assume that all students have the same initial choice set (the complete set of subject fields), that they apply for their preferred subset of fields and/or institutions and receive offers by one or more universities. Then they make their final decision based on this restricted set of offers. Hence, although the final choice is the student's one, the process leading to it is complex and is the result of the interactions between students and universities. This should be kept in mind every time we talk of students' choices in this paper. However, we would like to add that there is some evidence suggesting that our analysis of students' enrolled subject is very close to one of students' subject choice. In a recent article Leslie (2003) uses Universities' College Admissions Service (UCAS) data to build an indicator of quality of subjects. The author uses UCAS data for 1996-2001 and observes that 'each applicant is permitted to make up to six applications (except in medicine, which is restricted to four). Usually these six applications are in a well-defined subject area, but they need not to be so.' (p. 330-331). Another possible criticism to our analysis is that it might confound the effect of social class on subject choice with that on the probability of receiving an offer and accepting it. However, as Leslie (2003) observes, entry qualifications, especially A-levels score, are the key determinant of applicants' success and no other family or social background effects emerge. Therefore, on the basis of this evidence we can argue that in our model we are mainly estimating the effect of social class on applications rather than on offers and acceptances.

It must be noted that since the USR gathers information on university students only, all the empirical analysis that follows is conditional on enrollment in HE[7] and seeks to answer the following question: although there are social class differences in access to HE, once individuals from different social classes decide to enter HE, do they enroll in similar subjects?

4.2 Sample definition and descriptive statistics

From the cohort of students in each year 1981-1991, we select only non-mature students (students less than 21 when they entered HE), studying full time for a

degree qualification and we exclude overseas and married students. Moreover, since in our specification we want to control for the type and the level of performance at secondary school, we consider only students with A-level qualifications.[8] For the definition of the subject groups we take into account the predictions of the different hypotheses outlined in section 1 and the need to keep the econometric model estimable,[9] and aggregate all subjects in the following three broad subject areas:

1. 'Non-quantitative subjects' (abbreviated as NQS hereafter): Social Studies (excluding Economics), Mass Communications and Documentation, Languages and Related, Humanities, Creative Arts, Education, Combined degrees not included in the following category;

2. 'Quantitative subjects' (abbreviated as QS hereafter): Biological Sciences, Agriculture and Related, Physical Sciences, Mathematical Sciences and Informatics, Engineering and Technology, Architecture, Building and Planning, Economics, Business and Administration Studies, General Sciences Combined degrees;

3. Law and Medicine (L&M hereafter): Law, Medicine and Dentistry, Subjects Allied to Medicine.

In the period 1981-1991, the number of students satisfying our sample selection criteria rose by about 19%, from 48,024 to 57,096 units. However, the rise has been unevenly distributed across social classes, with students from social classes I and V,[10] for instance, rising only by 3.5% and decreasing by 1.2%, respectively, and those from social classes II, IIINM, IIIM and IV, rising by 25.2, 39.1, 9.15 and 52.3 percentage points,[11] respectively. The increase in the number of students was more equally distributed across subject groups. Both QS and NQS experienced an increase of around 20 percentage points, while the increase in L&M was about 7 percentage points lower.[12]

Table 1 reports some descriptive statistics on mean A-level score of entrant students by subject,[13] which can be interpreted as a raw measure of subject selectivity. L&M always ranked first in terms of mean A-level score of the student

intake, while the second place was occupied by quantitative subjects until 1986 and by non-quantitative subjects from 1987 onwards.

Table 2 shows average gross weekly occupational earnings of cohorts of student leavers since 1985.[14] For the whole period L&M ranked first in terms of average earnings, followed by quantitative and non-quantitative subjects, respectively. The coefficients of variation of earnings by study field generally show in the period of study a lower variation within QS, and a similar amount of earnings variation within the other subjects.

4.3 Choice of explanatory variables

The primary focus of this paper is on the effect of social class on the choice of undergraduate degree subject. Previous research has identified secondary school curriculum (Polachek, 1978), gender (Polacheck, 1978, Blakemore and Low, 1984) and forward-looking factors (Berger, 1988, Rochat and Demeulemeester, 2001, Montmarquette *et al.*, 2002) as the main determinants of undergraduate field of study.

In the present paper, we estimate a 'value added specification' of the subject choice model (see Todd and Wolpin, 2003) and do not consider the effect of forward-looking factors, such as expected incomes and academic performance. We decide to do so for several reasons. Firstly, past research has shown that the expected life-time flow of earnings is much more important than starting earnings for students' subject choice (see Berger, 1988) and that this flow is highly uncertain to students (see Wolter and Zbinden, 2002), while from USR data it is possible to have information on graduates' early occupational earnings only (i.e. six months after graduation, using the First Destination Supplement). In the absence of data on subjective earnings expectations by students, the construction of life-time subject specific expected earnings would require a substantial amount of discretion and assumptions on the part of the researcher.[15] Given that the inclusion of expected academic performance raises similar problems, the latter is also excluded from the present analysis. It must also be noted that our model is a reduced form model and therefore of descriptive nature. We want only to investigate whether there are statistically significant social class differences in the probability of enrolling in different subjects. Then, these differences may be originated by very different

factors such as differences in preferences or in the expected economic returns and costs of enrolling in different subjects, but unfortunately we do not have enough information to identify the various channels through which social class may exert its influence.

In detail, we include among the explanatory variables: gender, age at enrollment, secondary school type (not known, grammar, independent, comprehensive, 6th form college, other type), score in A-levels in specific subjects (biology, chemistry, economics, English, French, general studies, geography, history, mathematics, physics),[16] number of A-levels, best three A-level passes score, region of residence prior to university enrollment (inner London, outer London, other England, Scotland, Wales, Northern Ireland) and social class (I, II, IIINM, IIIM, IV, V, armed forces, non-workers, inadequately described).[17]

4.4 Models' fit

Table 3 reports some statistics for the TNP models estimated for the period 1981-1991. For all years, the Wald test for the null hypothesis that the coefficients on all regressors but the constant are equal to zero is strongly rejected. The Wald tests for the omission of the variables related to social class and pre-university school curriculum show that both sets of variables cannot be omitted from the model. However, it is the latter group of variables which accounts for most of the explanatory power of the model as the pseudo R^2 of the models with and without pre-university school variables show. Last but not least, in all years the estimated correlation ρ_{12} between the error components of the utilities (see section 3) of the QS and L&M groups is significantly different from zero, suggesting that the TNP model has to be preferred to a MNL model.

5 Empirical results on the effect of social class

In this section we comment on the estimated probabilities obtained from the TNP model.

The predicted probabilities of enrolling in the three different subject areas by social class are reported for each year in Table 4 and are computed as the means of

the individual predicted probabilities. Standard errors and 95% confidence intervals are computed using the delta method and Z critical values. We focus here only on the differences by social class.

Firstly, although Table 3 shows that social class variables are jointly highly significant, we observe in Table 4 that the predicted probabilities of enrolling in the different subject groups are generally not statistically different across social classes (i.e. confidence intervals overlap). Apart from statistical differences, we observe that some predictions of the theory are met by the data. In all years considered, except 1982, individuals from social class I had the smallest probability to enroll in QS (cf. hypothesis I in section 1). By contrast, the same individuals had in all years but 1991, the highest probability to enroll in Law and Medicine (cf. hypothesis IV in section 1). However, the differences are very low in magnitude and this influences their statistical significance. Moreover, there does not appear to be any dramatic change in the likelihood of enrolling in the different subject groups in the period under study.

Thus, our analysis appears to show that the structure of the UK higher education system during the period 1981-1991 was able to ensure that individuals with different social backgrounds had equal opportunities of accessing different subjects at tertiary level (conditional on accessing HE). Recall that the system was characterized by the absence of undergraduate tuition fees and by the provision of means-tested maintenance grants for economically disadvantaged students. Both these features of the UK higher education system were likely to attenuate differences across social classes in the probability of enrolling in the different subjects by making the pecuniary costs of enrolling in the different fields very similar across social classes and subjects. Moreover, in a period in which the number of graduates was not very high, possessing a degree was probably sufficient per-se to ensure high earnings in the labour market and more important than social class origin.[18] On the basis of these results it might be interesting to replicate our analysis for more recent years, since the 1990s were characterized by a gradual substitution of the maintenance grants with repayable loans. As stated by Callender (2003) the replacement of student maintenance grants with subsidised loans marked a switch from a system granting a large subsidy to lower income students to a less generous

system benefiting all students. Although student loans might be very close to implementing perfect capital markets, since students can borrow against their future incomes at a zero real interest rate, they are surely less generous than maintenance grants for low income students. In particular, their introduction may have had some consequences not only on student access to HE but also in terms of differentiating students' choices across social classes. Low income students, who are more risk adverse, might have preferred to enroll in less selective and "easier" subjects, in which the probability of failure is lower or that of achieving a 'good' (first or upper second) degree class higher, given the growing importance of degree class over time (see Naylor *et al.* 2003), or to enroll in the subjects in which there is less earnings dispersion. Evidence in this direction is implicitly provided by Metcalf (2005) who finds that the introduction of home students tuition fees in 1998 increased student debt (circa the same amount of fees) in particular for students who did not receive financial support from their parents, and term time employment for the same category of students. Moreover, term time employment was more popular among students in less demanding courses, such as social sciences and humanities. As to student's satisfaction with respect to the university experience, student debt was the primarily cited cause of regret. Overall this evidence suggests that low income students might decide to enroll in the less time demanding courses in order to be able to work during term time and also accumulate more debt and be dissatisfied with their university education after getting a degree. Moreover, Pitcher and Purcell (1998) using data from the 'Great Expectation' survey of UK final year undergraduates in 1996 found that three quarters of students in their sample accumulated a debt of more than £ 500 while studying and that about one third of this group felt that they would have had to take any kind of employment in order to pay their debts with negative consequences for their careers.[19] This problem might have become worse after the introduction of student fees in 1998. And, of course, it would also be interesting to study the consequences of imposing top-up fees and of differentiating the fees by subject, which have the potential of producing further unequalising effects on the choices of students from different social classes.

The bulk of explanatory power of our model of subject choice can be ascribed to the type of pre-university school curriculum and performance, proxied by the A-

levels score, number of A-levels, school type and type of A-levels with the relative performance. The drop in the pseudo R^2 when these variables are omitted (see Table 3) shows that they have a high explanatory power over and above social class (as the models with full controls show), which in turn has only a limited influence on pre-university school curriculum. Indeed, the explanatory power of social class remains low also when secondary school variables, on which the former may have an influence, are not controlled for. This confirms the findings by van de Werfhorst et al. (2003) and is relatively good news in terms of intergenerational mobility and equal educational opportunities of the UK university system in the 1980s as far as subject choice is concerned: social class did not appear to be the main determinant of students' differences of undergraduate subject studied. Our analysis appears to implicitly attribute a major role to other individual or family characteristics, such as parenting quality or students' ability and motivation, or to quality of schools and teachers, which affect the type of secondary school curriculum and A-level performance. However, we do not exclude that there might have been other forms of educational inequalities across social classes, for instance, in terms of the type of institution enrolled (polytechnics vs universities, Russel group institution vs other universities).[20]

5.1 Sensitivity analysis

We have seen that our econometric analysis excludes the presence of statistically significant differences across individuals from different social classes on the probability of enrolling in different subject groups. However, we may wonder whether there are differences within broadly defined subject groups, and whether the results are determined by the specific econometric model chosen.

We address both issues by considering a more detailed definition of subject groups. In particular: Medicine, Law, Sciences, Technical, Economics and Business, Mathematics, Soft Social Sciences, Art and Humanities.[21]. Since the high number of subject groups considered does not allow the estimation of a multinomial probit model, we use instead a flexible-thresholds ordered probit model (see the Appendix). However, estimation of such a model requires an ordering of the discrete dependent variable. We ranked the subjects in ascending order of occu-

pational weekly earnings. In particular, we matched 1985-1991 university students leavers data from the USR with occupational earnings from the NES survey and computed an average of the occupational earnings by subject group. On the basis of the average weekly occupational earnings in each year the subjects were ordered to estimate the flexible-thresholds ordered probit model. The ranking of the subjects is shown in Table 5 and is very similar across years. We report in Tables 6-7 the estimates for a benchmark year, 1985.[22] Concerning 1985, it must be noted that the ordering by weekly occupational earnings of the broader subject groups shown in Table 2 is generally preserved (i.e. Law and Medicine graduates had higher earnings than all QS graduates who in turn earned more than NQS graduates), the sole exception being Sciences, whose graduates were at the bottom of the earnings distribution. Table 6 shows that by considering a finer disaggregation of subject groups, imposing an ordering and using a different econometric model our results do not change: differences across social classes are never statistically significant. Moreover, Table 6 shows that when we compute the aggregate probabilities of the broader subjects considered in the previous section (QS, NQS, Law and Medicine) by summing the probabilities of the finer subjects, they are very similar to those obtained from the TNP model. In summary, our results do not appear to be driven by the aggregation of subjects or by the type of econometric model chosen.

6 Concluding remarks

In this paper we investigated social class influences on degree subject choice at the undergraduate level, conditional on enrolling in HE in the UK. We have used data for several cohorts of university students in the 1980s and early 1990s to estimate a trinomial probit model of subject choice (Quantitative Subjects, Non-Quantitative Subjects, Law and Medicine). We have also considered a finer disaggregation of subjects and an alternative econometric model (a flexible-thresholds ordered probit) and showed the robustness of our results. From our empirical analysis:

1. We do not find statistically significant differences among social classes in the probability of enrolling in different subjects in the period 1981-1991.

2. In the 'value added specification' of our model (i.e., controlling for secondary school variables, see Todd and Wolpin, 2003) social class explains only a small part of the variation in subject enrolled, while secondary school curriculum (school type, A-level score and number and performance in specific A-levels) has a much higher explanatory power. The explanatory power of social class remains low also in a 'contemporaneous specification' of the model of subject enrolled, where secondary school variables are omitted, suggesting that its role may be limited also at early stages of the educational process.

Both findings can be interpreted as good news in terms of intergenerational mobility of the UK university system during the 1980s as far as subject choice is concerned, in the sense that a student's choice of study field was made on the basis of characteristics generally unrelated to social class.

For future research, it would be interesting to replicate the analysis in this paper for more recent cohorts of university students using HESA data. Indeed, the recent changes in the UK university system, such as the gradual replacement of student maintenance grants with student loans and the introduction of tuition fees in 1998 may have contributed to differentiating the degree subject choices of students with different social class backgrounds and give some useful insights into the potential effects of introducing top-up fees and fees differentiated by subject from 2006.

Notes

[1]For some studies related to intergenerational mobility in Britain see Dearden *et al.* (1997), Blanden *et al.* (2003) and Machin and Gregg (2003), among others. Machin and Gregg (2003) observe, for instance, that the educational expansion of the late 1980s early 1990s benefited especially high social class students, contributing to a decrease in intergenerational mobility.

[2]Because they need to borrow and pay back student loans while high social class students usually have access to cheaper, even free, family resources to finance higher education.

[3]Using Universities' Statistical Record data for the UK, Mancini (2003) finds, for instance, that working class graduates in law are significantly more likely to be unemployed than their wealthier peers.

[4]See Blanden and Machin (2004) for an outline of recent changes in the UK system of Higher Education. In brief, the expansion in the UK education system was partly implemented by reducing the generosity in student support. In particular, the major changes were the freezing of maintenance grants in 1990 and their progressive replacement with subsidised loans; the introduction of home students undergraduate fees of 1,000 pounds per year, the increase in the maximum loan and the introduction of an income-contingent repayment system following the 1997 Dearing Report. Beginning from 2006, universities that meet some requirements will be able to charge top-up fees which may reach a maximum of 3,000 pounds.

[5]The assumption that also the variance of the second error term is one is not strictly necessary, but is often found in empirical applications since it helps the model estimation. In this case we assume that the error terms in the equations 5 and 6 are standard normal.

[6]For a review on the multinomial probit model see also Weeks (1997).

[7]Such as all the literature reported in section 2.

[8]Since we want to focus on the choice of a typical student and investigate the effect of his/her parental background, the decision to restrict the analysis to non-mature students, studying full-time, i.e. individuals for which study is the main activity and who are likely to be more affected by their parental and social background is natural. Mature students may have enrolled in HE after working for a period and have accumulated the financial resources necessary to enroll in HE, in any case they are likely to be more independent of their families. Moreover, the USR data do not provide family background information for mature students. We exclude students with non-traditional entry qualifications into

HE since the level of secondary school performance, which we consider as a control for students' ability, is not available. However, A-level entrants represent the vast majority of university students in the period studied (1981-1991).

[9]See section 3.

[10]Some abbreviations commonly used in the UK for social classes are: I (professional), II (intermediate), IIINM (skilled non manual), IIIM (skilled manual), IV (partly skilled), V (unskilled). In the USR data social class was built using parents' occupation. We are grateful to Abigail McNight for providing us with the mapping information.

[11]The high increase in social class IV is, however, partly determined by the low initial number of students with this social background in 1981.

[12]Although the figure for this group reflects the slower dynamic for Medicine, since the number of medical students is determined by the Government.

[13]A-level scores are computed according to the UCAS method: A=10, B=8, C=6, D=4, E=2.

[14]The author wishes to thank Abigail McNight, Robin Naylor and Jeremy Smith for providing data on earnings. Weekly occupational earnings are obtained by matching First Destination Supplement data, in the USR, with data from the New Earnings Survey. See Bratti and Mancini (2003) for a more detailed description of the matching procedure.

[15]Some studies, such as Dominitz and Manski (1996), Betts (1996), and Brunello *et al.* (2001), cast doubts on students' ability to predict their life-time future earnings. These studies generally show a large heterogeneity in students' beliefs about current earnings, which reflects a large variation in students' information. Finally, as noticed by Dominitz and Manski (1996): 'incorrect assumptions can yield incorrect inferences about the way students make schooling decisions' (p. 3).

[16]In order to keep the model tractable we choose to include only the most popular subjects for which a score as well as a pass indicator is available.

[17]For the explanation of the abbreviations for social classes see footnote 10.

[18]Compared to a situation in which the supply of graduates is high and 'family networks' may be important for employers to screen among graduates.

[19]Stewart and Swaffield (1999), for instance, using UK data find that the probability of being low paid depends on low pay in the previous year.

[20]This may happen since there are positive economic returns to attending prestigious Universities as shown by Chevalier and Conlon (2003).

[21]The composition is: Medicine, Law, Sciences (Biological Sciences, Agriculture, Physical Sciences, General Sciences Combined degrees), Technical (Computing, Engineering,

Technology, Architecture), Economics and Business, Mathematics, Soft Social Sciences (Social studies excluding economics, Politics, Mass Communications, General Social Sciences Combined degrees), Art and Humanities (Classics and Literature, Modern Euro Languages, Other Languages, Humanities, Creative Arts, Education, Other combined degrees).

[22]The same sensitivity analysis was also replicated for 1991 and showed the same results.

Appendix. The flexible-thresholds ordered probit model

In this section, we offer an economic rationalization for the flexible-thresholds ordered probit (FT-OP, hereafter) model, introduced by Pradhan and van Soest (1995), in the spirit of Cameron and Heckman (1998). An individual has to choose an undergraduate subject among a group $j = 1, ..., J$ of possible alternatives. Given the individual's characteristics x_i the cost (both pecuniary and non-pecuniary) of enrolling subject j is $c(j|x_i)$ while the benefit of enrolling in the same subject is $R(j|x_i, \epsilon_i)$ where ϵ_i is a person specific shifter of the return to field j. The subjects are such that high return study fields also imply higher costs, e.g. the highly rewarded subjects are also the most selective. Therefore both $c(j|x_i)$ and $R(j|x_i)$ are increasing in j (i.e. subjects are ordered in increasing order of returns and costs). We define the utility for individual i in subject j, i.e. V_{ij}, as the difference between the return and the cost of acquiring education in field j. The optimal field of study is determined for each individual by solving the problem:

$$\underset{j}{Max}[R(j|x_i, \epsilon_i) - c(j|x_i)] \tag{10}$$

where $j = 1, ..., J$, and J is the field with the highest economic return (which does not necessarily imply that it is also the highest utility field for individual i).

Let us assume that ϵ_i is stochastic and continuously distributed, such that $\epsilon_i \perp x_i$ and:

$$c(s|x_i) = \sum_{j=1}^{s} \exp(\delta_1) \prod_{z=2}^{j} \exp(\delta_j(x_i)) \tag{11}$$

while

$$R(s|x_i, \epsilon_i) = \epsilon_i \sum_{j=1}^{s} \exp(\phi_1)\psi(x_i) \prod_{z=2}^{j} \exp(\phi_j(x_i)). \tag{12}$$

Hence, in this case both cost and return depend on individuals' characteristics. We assume that students enroll in the field which maximize their utility. If $1 < s < J$ is the optimal undergraduate field for individual i, then:

$$\frac{\exp(\delta_1)\prod_{j=2}^{s}\exp(\delta_j(x_i))}{\exp(\phi_1)\prod_{j=2}^{s}\exp(\phi_j(x_i))\psi(x)} \leq \epsilon_i \leq \frac{\exp(\delta_1)\prod_{j=2}^{s+1}\exp(\delta_j(x_i))}{\exp(\phi_1)\prod_{j=2}^{s+1}\exp(\phi_j(x_i))\psi(x)}. \tag{13}$$

By further assuming that $\phi_j(x_i) = \exp(\phi_j' x_i)$ and $\delta_j(x_i) = \exp(\delta_j' x_i)$ and $\psi(x_i) = \exp(x_i\beta)$ the last expression can be rewritten as:

$$\frac{\exp(\delta_1)\prod_{j=2}^{s}\exp(\exp(\delta_j' x_i))}{\exp(\phi_1)\exp(x_i\beta)\prod_{j=2}^{s}\exp(\exp(\phi_j' x_i))} \leq \epsilon_i \leq \frac{\exp(\delta_1)\prod_{j=2}^{s+1}\exp(\exp(\delta_j' x_i))}{\exp(\phi_1)\exp(x_i\beta)\prod_{j=2}^{s+1}\exp(\exp(\phi_j' x_i))}. \tag{14}$$

If we define $\exp(l_s(x_i)) \equiv \exp(\delta_1 - \phi_1)\prod_{j=2}^{s}\exp((\delta_j - \phi_j)' x_i)$, $\delta_j - \phi_j \equiv \lambda_j$ and assume that ϵ_i is standard log-normally distributed,

$$Pr(j = s|X = x_i) = \Phi\big[\lambda_1 + \sum_{j=2}^{s}\exp(\lambda_j' x_i) - \beta' x_i \leq \mu_i \leq \lambda_1 + \sum_{j=2}^{s+1}\exp(\lambda_j' x_i) - \beta' x_i\big]. \tag{15}$$

where $\mu_i = ln(\epsilon_i)$ and $\Phi(\cdot)$ is the standard normal distribution function, we obtain the flexible-thresholds ordered probit model introduced by Pradhan and van Soest (1995). The thresholds $l_s(x_i)$ are allowed to depend on individual characteristics x_i. As Pradhan and van Soest (1995) observed, this model allows greater flexibility compared to the standard ordered probit model. Indeed, model identification only requires one threshold to be fixed. Therefore the generality of the model can be increased by letting the other thresholds depend on individual characteristics. In particular, while the choice of the lowest return field depends on the index $\beta' x_i$ only, the choices of the other subjects also depend on the $\lambda_j' x_i$'s indexes. Despite being more flexible than the ordered probit model the flexible-thresholds ordered probit model requires the ordering of the outcome variable and it is therefore less general than a multinomial probit model although much easier to estimate.

References

Becker, G. (1975) *Human capital*. Chicago: Chicago University Press.

Betts, J. R. (1996), 'What do students know about wages? Evidence from a survey of undergraduates', *Journal of Human Resources*, 31: 27-56.

Berger, M. C. (1988), 'Predicted future earnings and choice of college major', *Industrial and Labor Relations Review*, 41: 418-429.

Blackaby, D. H., Murphy, P. D., and O'Leary, N. C. (1999), 'Graduate earnings in Britain: A matter of degree?', *Applied Economics Letters*, 6: 311-315.

Blanden, J., Goodman, A., Gregg, P. and Machin, S. (2003), 'Changes in intergenerational mobility in Britain', CEPR Discussion Paper n. 517, LSE, London.

Blanden, J., Machin, S. (2004), 'Educational inequality and the expansion of UK higher education', *Scottish Journal of Political Economy*, 51: 230-249.

Boudon, R. (1974), *Education, opportunity and social inequality*. New York: Wiley.

Bourdieu, P. (1984), *Distinction: A social critique of the judgement of taste*. London: Routledge and Kegan Paul.

Bratti, M. and Mancini, L. (2003), 'Differences in early occupational earnings of UK male graduates by degree subject: Evidence from the USR, 1980-1993', IZA Discussion Paper No. 890, IZA, Bonn.

Bratti, M., McKnight, A., Naylor, R. and Smith, J. (2004), 'Higher education outcomes, graduate employment and university performance indicators" *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 167: 475-496.

Brunello, G., Lucifora, C. and Winter-Ebmer, R. (2001), 'The wage expectations of European college students', IZA Discussion Paper No. 299, IZA, Bonn, forthcoming in *Journal of Human Resources*.

Bunch, D. S. (1991), 'Estimability in the multinomial probit model', *Transportation Research B*, 25: 1-12.

Callender, C. (2003), 'Student financial support in higher education: Access and exclusion. In M. Tight (ed.), *Access and exclusion: international perspectives on higher education research*. London: Elsevier Science.

Cameron, S. and Heckman, J. (1998) 'Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of American males', *Journal of Political Economy*, 106: 262-333.

Chevalier, A. and Conlon, G. (2003), 'Does it pay to attend a prestigious university?', IZA Discussion Paper No. 848, IZA, Bonn.

Chevalier, A., Conlon, G., Galindo-Rueda, F. and Mcnally, S. (2002), 'The returns to higher education teaching', Report to the Department of Education and Skills, Center for the Economics of Education, LSE, London.

Chevalier, A. and Lanot, G. (2002), 'The relative effect of family characteristics and financial situation on educational achievement', *Education Economics*, 10: 165-182.

Dansie, B. (1985), 'Parameter estimability in the multinomial probit model', *Transportation Research B*, 19: 526-528.

Davies, S. and Guppy, N. (1997), 'Fields of study, college selectivity, and student inequalities in higher education', *Social Forces*, 75: 1417-1438.

Dearden, L., Machin, S. and Reed, L. (1997), 'Intergenerational mobility in Britain', *Economic Journal*, 107: 47-64.

Doeringer, P. and Piore, M. (1971), *Internal labor markets and manpower analysis*. Lexington, MA: D.C. Health.

Dominitz, J., and Manski, C. F. (1996), 'Eliciting student expectations of the returns to schooling', *Journal of Human Resources* 31: 1-26.

Ermisch, J. and Francesconi, M. (2001a) 'Family matters: Impacts of family background on educational attainments', *Economica*, 68: 137-156.

Heckman, J., and Sedlacek, G. (1985), 'Heterogeneity, aggregation, and market wage functions: An empirical model of self selection in the labor market', *Journal of Political Economy*, 93: 1077-1125.

Keane, M. P. (1992), 'A note on identification in the multinomial probit model', *Journal of Business and Economics Statistics*, 10: 193-200.

Kelsall, R. K., Poole, A. and Kuhn, A. (1972), *Graduates: The sociology of an elite*. London: Methuen.

Leslie, D. (2003) 'Using the success to measure quality in British higher education: Which subjects attract the best-qualified students?, *Journal of The Royal Statistical Society, Series A*, 166: 329-347.

Machin, S. and Gregg, P. (2003), 'A lesson for education: University expansion and falling income mobility', *New Economy* 10: 194-198.

Mancini, L. (2003), *Higher education in the UK and the market for labour: Evidence from the Universities' Statistical Record*, Ph.D. Thesis, University of Warwick, Coventry, December 2003.

Metcalf, H. (2005), 'Paying for university: The impact of increasing costs on student employment, debt and satisfaction', *National Institute Economic Review*, 191: 106-117.

Micklewright, J. (1989) 'Choice at sixteen', *Economica*, 56: 25-39.

Montgomery, J. D. (1991), 'Social networks and labor-market outcomes: Toward an economic analysis', *American Economic Review*, 81: 1408-1418.

Montmarquette, C., Cannings, K. and Mahseredjian, S. (2002), 'How do young people choose college majors?', *Economics of Education Review*, 21: 543-556.

Naylor, R., Smith, J. and McKnight, A. (2003), 'Returns to educational performance: evidence from UK graduates' first destination labour market outcomes", mimeo, University of Warwick, Coventry.

Oosterbeek, H. and Webbink, D. (1997), 'Is there a hidden technical potential?', *De Economist*, 145: 159-177.

Pitcher, P. and Purcell, J. (1998), 'Diverse expectations and access to opportunities: Is there a graduate labour market?', *Higher Education Quarterly*, 52: 179-203.

Pradhan, M. and van Soest, A. (1995), 'Formal and informal sector employment in urban areas of Bolivia', *Labour Economics*, 2: 275-297.

Rice, P. G. (1987) 'The demand for post-compulsory education in the UK and the effects of educational maintenance allowances', *Economica*, 54: 465-75.

Rochat, D. and Demeulemeester, J.-L. (2001), 'Rational choice under unequal constraints: The example of Belgian higher education', *Economics of Education Review*, 20: 15-26.

Smith, J., McKnight, A. and Naylor, R. A. (2000), 'Graduate employability: Policy and performance in higher education in the UK, *Economic Journal*, 110: F382-411.

Stewart, M. B. and Swaffield, J. K. (1999), 'Low pay dynamics and transition probabilities', *Economica*, 66: 23-42.

Todd, P. E. and Wolpin, K. I. (2003), 'On the specification and estimation of the production function for cognitive achievement', *Economic Journal*, 113: F3-F33.

van de Werfhorst, H. G., de Graaf, N. D. and Kraaykamp, G. (2001), 'Intergenerational resemblence in field of study in the Netherlands', *European Sociological Review*, 17: 275-293.

van de Werfhorst, H. G., Sullivan, A. and Cheung, S. Y., (2003), 'Social class, ability and choice of subject in secondary and tertiary education in Britain', *British Educational Research Journal*, 29: 41-62.

Walker, I. and Zhu, Y. (2001), *The returns to education Evidence from the Labour Force Surveys.* Department for Education and Skills, Research Report 313, DfES, London.

Wolter, S. C., and Zbinden, A. (2002), 'Labour market expectations of Swiss university students', *International Journal of Manpower*, 23: 458-470.

Weeks, M. (1997), 'The multinomial probit model revisited: A discussion of parameter estimability, identification and specification testing', *Journal of Economic Surveys*, 11: 297-320.

Tables

Table 1: Average A-level score of entrant students by subject (USR data)

Year	QS		NQS		L&M	
	mean	c.v.	mean	c.v.	mean	c.v.
1981	20.418	0.308	20.220	0.292	23.774	0.195
1982	21.130	0.281	21.071	0.264	24.214	0.179
1983	22.020	0.247	21.699	0.237	24.663	0.161
1984	22.123	0.250	21.849	0.237	25.027	0.155
1985	22.076	0.255	21.657	0.239	25.055	0.157
1986	21.947	0.259	21.774	0.234	25.140	0.153
1987	21.728	0.266	21.897	0.231	24.910	0.164
1988	21.794	0.268	21.971	0.231	25.015	0.161
1989	21.883	0.270	22.324	0.214	25.199	0.158
1990	21.939	0.269	22.631	0.208	25.097	0.164
1991	21.710	0.277	23.129	0.197	25.110	0.172

Note. QS: 'Quantitative Subjects'; NQS: 'Non-Quantitative Subjects'; L&M: Law and Medicine (see section 4.2). The table reports average A-level scores of entrant students by subject and the coefficient of variation (c.v.) within subject groups. The average A-level score are computed on the estimation samples (i.e. non-mature, non-overseas, unmarried students studying for a first degree qualification and with A-levels).

Table 2: Average gross weekly occupational earnings of leaving students by subject - 3 groups

Year	QS		NQS		L&M	
	mean	c.v.	mean	c.v.	mean	c.v.
1985	445.5	0.2	430.5	0.2	564.6	0.2
1986	487.1	0.2	468.0	0.2	612.1	0.2
1987	537.4	0.2	524.8	0.2	699.8	0.2
1988	598.7	0.2	567.9	0.2	785.6	0.2
1989	656.6	0.2	626.9	0.2	871.6	0.2
1990	677.1	0.2	661.3	0.2	980.9	0.2
1991	710.7	0.2	690.8	0.3	1044.7	0.3
1992	735.7	0.2	716.2	0.3	1071.4	0.3
1993	761.4	0.2	747.1	0.3	1085.3	0.3

Note. QS: 'Quantitative Subjects'; NQS: 'Non-Quantitative Subjects'; L&M: Law and Medicine (see section 4.2). The table reports average gross weekly occupational earnings in current pounds (from the New Earnings Survey) of leaving students by subject and the coefficient of variation (c.v.) within subjects. Average earnings are computed on the estimation samples (i.e. non-mature, non-overseas, unmarried students studying for a first degree qualification and with A-levels).

Table 3: Model diagnostics

Statistics	Years						
	1981	1982	1983	1984	1985	1986	1987
Log-likelihood	-27,978.0	-26,887.6	-26,231.0	-26,826.2	-27,787.3	-27,678.0	-28,038.7
Wald test: overall significance$^{(a)}$	13,752.1	13,718.4	14,206.0	14,677.4	16,015.6	16,027.0	15,872.1
(p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Wald test: social class$^{(b)}$	80.80	203.98	97.76	52.66	98.29	113.92	95.19
(p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Wald test: pre-university variables$^{(c)}$	12,224.1	12,455.7	13,035.9	13,357.4	14,645.2	14,590.8	14,424.0
(p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Wald test: school type$^{(d)}$	120.3	88.6	46.8	88.8	78.8	113.3	110.4
(p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Pseudo R^2:$^{(e)}$							
i. model with secondary education	0.41	0.41	0.41	0.41	0.41	0.40	0.40
ii. model withouth secondary education	0.05	0.05	0.04	0.04	0.05	0.04	0.04
$\rho_{12}$$^{(f)}$	0.70	0.73	0.74	0.70	0.63	0.61	0.61
(s.e.)	0.03	0.03	0.03	0.03	0.03	0.03	0.03
N. observations	48,024	45,837	44,412	45,336	47,601	46,568	46,914

continued

Statistics	Years			
	1988	1989	1990	1991
Log-likelihood	-28,905.1	-31,239.1	-33,443.6	-34,060.5
Wald test: overall significance[(a)]	17,243.3	19,522.9	20,236.2	22,006.7
(p-value)	0.00	0.00	0.00	0.00
Wald test: social class[(b)]	86.91	77.74	98.75	88.28
(p-value)	0.00	0.00	0.00	0.00
Wald test: pre-university variables[(c)]	15,668.5	17,767.9	18,547.6	20,280.8
(p-value)	0.00	0.00	0.00	0.00
Wald test: school type[(d)]	116.9	131.7	175.6	166.0
(p-value)	0.00	0.00	0.00	0.00
Pseudo R^{2}:[(e)]				
i. model with secondary education	0.39	0.40	0.39	0.39
ii. model withouth secondary education	0.04	0.04	0.04	0.05
ρ_{12}[(f)]	0.56	0.48	0.41	0.39
(s.e.)	0.04	0.04	0.04	0.05
N. observations	48,303	52,871	55,782	57,096

Note. [(a)] $\text{Chi}^2(64)$; [(b)] $\text{Chi}^2(16)$; [(c)] $\text{Chi}^2(24)$. Pre-university variables include A-level number, A-level score (UCAS formula) and A-level performance by subject; [(d)] $\text{Chi}^2(10)$; [(e)] The Pseudo R^2 is computed as: 1-(loglikelihood only-constant model/loglikelihood full model); [(f)] Coefficient of correlation between the alternatives Law and Medicine and Quantitative Subjects.

Table 4: Predicted probabilities (%), with standard errors and 95% confidence intervals, of enrolling in the different study fields by social class

Social Class	Subject group								
		QS			NQS			L&M	
	Prob	left	right	Prob	left	right	Prob	left	right
1981									
I	0.48	0.46	0.51	0.36	0.34	0.38	0.16	0.13	0.18
II	0.50	0.48	0.52	0.37	0.36	0.39	0.13	0.11	0.15
IIINM	0.49	0.47	0.51	0.37	0.36	0.39	0.13	0.11	0.15
IIIM	0.50	0.48	0.53	0.37	0.36	0.39	0.12	0.10	0.14
IV	0.49	0.46	0.51	0.39	0.37	0.41	0.12	0.10	0.15
V	0.52	0.48	0.57	0.36	0.33	0.40	0.12	0.08	0.15
Armed forces	0.48	0.44	0.52	0.40	0.37	0.43	0.12	0.09	0.15
Inadequately described	0.50	0.47	0.53	0.37	0.35	0.40	0.12	0.10	0.15
Non workers	0.49	0.45	0.52	0.39	0.36	0.41	0.12	0.10	0.15
1982									
I	0.47	0.45	0.49	0.36	0.34	0.38	0.17	0.14	0.19
II	0.49	0.47	0.51	0.37	0.35	0.39	0.14	0.12	0.16
IIINM	0.51	0.48	0.53	0.36	0.34	0.38	0.13	0.11	0.15
IIIM	0.51	0.49	0.54	0.36	0.35	0.38	0.12	0.10	0.14
IV	0.50	0.47	0.53	0.36	0.34	0.38	0.14	0.11	0.16
V	0.47	0.43	0.52	0.36	0.33	0.40	0.16	0.12	0.21
Armed forces	0.49	0.45	0.53	0.37	0.34	0.40	0.14	0.11	0.17
Inadequately described	0.49	0.46	0.52	0.38	0.36	0.41	0.13	0.10	0.15
Non workers	0.37	0.34	0.41	0.37	0.34	0.39	0.26	0.23	0.29
1983									
I	0.46	0.44	0.49	0.36	0.34	0.38	0.18	0.16	0.20
II	0.49	0.47	0.52	0.36	0.35	0.38	0.15	0.12	0.17
IIINM	0.50	0.47	0.52	0.36	0.34	0.37	0.15	0.12	0.17
IIIM	0.51	0.48	0.53	0.36	0.34	0.38	0.13	0.11	0.16
IV	0.51	0.48	0.53	0.36	0.34	0.38	0.14	0.11	0.16
V	0.52	0.48	0.57	0.38	0.34	0.41	0.10	0.07	0.14
Armed forces	0.47	0.43	0.51	0.38	0.34	0.41	0.15	0.12	0.19
Inadequately described	0.48	0.45	0.51	0.36	0.34	0.38	0.16	0.13	0.19
Non workers	0.48	0.45	0.52	0.37	0.35	0.40	0.14	0.12	0.17
1984									
I	0.47	0.45	0.50	0.36	0.34	0.37	0.17	0.15	0.19
II	0.49	0.47	0.51	0.36	0.35	0.38	0.15	0.13	0.16
IIINM	0.49	0.47	0.52	0.36	0.34	0.38	0.15	0.13	0.17
IIIM	0.50	0.48	0.52	0.36	0.34	0.38	0.14	0.12	0.16
IV	0.49	0.47	0.52	0.36	0.34	0.38	0.15	0.13	0.17
V	0.48	0.43	0.53	0.38	0.35	0.41	0.14	0.10	0.18
Armed forces	0.47	0.43	0.51	0.38	0.35	0.41	0.15	0.12	0.19
Inadequately described	0.48	0.45	0.51	0.36	0.34	0.39	0.15	0.13	0.18
Non workers	0.47	0.43	0.50	0.39	0.36	0.41	0.15	0.12	0.18
1985									
I	0.49	0.47	0.51	0.35	0.33	0.36	0.17	0.15	0.19
II	0.51	0.49	0.53	0.36	0.34	0.37	0.14	0.12	0.15
IIINM	0.51	0.49	0.54	0.35	0.33	0.37	0.13	0.11	0.15
IIIM	0.52	0.50	0.54	0.35	0.33	0.36	0.13	0.11	0.15
IV	0.51	0.48	0.53	0.35	0.33	0.37	0.14	0.12	0.16
V	0.52	0.48	0.56	0.35	0.32	0.39	0.12	0.09	0.15
Armed forces	0.49	0.45	0.53	0.38	0.35	0.41	0.12	0.09	0.15
Inadequately described	0.52	0.49	0.55	0.35	0.32	0.37	0.13	0.11	0.16
Non workers	0.49	0.46	0.52	0.38	0.35	0.40	0.13	0.11	0.16

continued

Social Class	Subject group QS			NQS			L&M		
	Prob	left	right	Prob	left	right	Prob	left	right
1986									
I	0.48	0.46	0.50	0.35	0.33	0.37	0.17	0.15	0.19
II	0.51	0.49	0.53	0.36	0.34	0.37	0.14	0.12	0.16
IIINM	0.51	0.49	0.54	0.35	0.33	0.37	0.14	0.12	0.16
IIIM	0.53	0.50	0.55	0.34	0.32	0.36	0.13	0.11	0.15
IV	0.52	0.49	0.54	0.34	0.32	0.36	0.14	0.12	0.16
V	0.46	0.42	0.50	0.37	0.33	0.40	0.17	0.13	0.21
Armed forces	0.52	0.48	0.56	0.36	0.33	0.39	0.12	0.09	0.15
Inadequately described	0.51	0.48	0.54	0.34	0.32	0.37	0.15	0.12	0.17
Non workers	0.50	0.46	0.53	0.36	0.33	0.39	0.14	0.12	0.17
1987									
I	0.48	0.46	0.50	0.35	0.33	0.37	0.17	0.15	0.19
II	0.51	0.49	0.53	0.35	0.33	0.37	0.14	0.12	0.16
IIINM	0.53	0.50	0.55	0.34	0.32	0.36	0.13	0.12	0.15
IIIM	0.52	0.50	0.55	0.34	0.32	0.35	0.14	0.12	0.16
IV	0.51	0.49	0.54	0.34	0.32	0.36	0.14	0.12	0.17
V	0.53	0.48	0.57	0.33	0.29	0.36	0.14	0.11	0.18
Armed forces	0.51	0.47	0.55	0.34	0.31	0.37	0.15	0.11	0.18
Inadequately described	0.51	0.48	0.54	0.34	0.32	0.36	0.15	0.13	0.18
Non workers	0.54	0.50	0.57	0.35	0.32	0.37	0.12	0.09	0.14
1989									
I	0.49	0.47	0.51	0.37	0.35	0.38	0.15	0.13	0.16
II	0.51	0.49	0.53	0.36	0.34	0.37	0.13	0.11	0.14
IIINM	0.53	0.51	0.55	0.35	0.34	0.37	0.12	0.10	0.14
IIIM	0.52	0.50	0.54	0.35	0.34	0.37	0.13	0.11	0.14
IV	0.51	0.48	0.53	0.35	0.34	0.37	0.14	0.12	0.16
V	0.52	0.47	0.56	0.34	0.31	0.38	0.14	0.11	0.18
Armed forces	0.51	0.47	0.55	0.38	0.34	0.41	0.12	0.09	0.14
Inadequately described	0.52	0.50	0.55	0.34	0.32	0.37	0.13	0.11	0.16
Non workers	0.53	0.50	0.56	0.36	0.33	0.38	0.12	0.09	0.14
1990									
I	0.48	0.46	0.50	0.38	0.36	0.39	0.15	0.13	0.16
II	0.50	0.48	0.52	0.38	0.36	0.39	0.12	0.11	0.14
IIINM	0.51	0.49	0.53	0.36	0.35	0.38	0.13	0.11	0.15
IIIM	0.51	0.49	0.53	0.36	0.35	0.38	0.13	0.11	0.15
IV	0.50	0.48	0.53	0.36	0.34	0.38	0.13	0.11	0.15
V	0.51	0.46	0.55	0.35	0.31	0.38	0.14	0.11	0.18
Armed forces	0.50	0.46	0.54	0.40	0.37	0.43	0.10	0.08	0.13
Inadequately described	0.49	0.46	0.52	0.36	0.33	0.38	0.15	0.13	0.18
Non workers	0.51	0.49	0.54	0.35	0.33	0.38	0.13	0.11	0.16
1991									
I	0.48	0.46	0.50	0.38	0.36	0.39	0.14	0.13	0.16
II	0.50	0.48	0.52	0.38	0.36	0.39	0.12	0.11	0.14
IIINM	0.51	0.49	0.53	0.36	0.35	0.38	0.12	0.11	0.14
IIIM	0.51	0.49	0.53	0.37	0.35	0.38	0.13	0.11	0.14
IV	0.51	0.48	0.53	0.36	0.34	0.38	0.14	0.12	0.16
V	0.50	0.46	0.55	0.34	0.31	0.38	0.16	0.12	0.20
Armed forces	0.50	0.46	0.54	0.38	0.35	0.41	0.12	0.09	0.15
Inadequately described	0.49	0.46	0.51	0.38	0.36	0.40	0.13	0.11	0.15
Non workers	0.52	0.48	0.55	0.35	0.32	0.37	0.14	0.11	0.17

Note. QS: 'Quantitative Subjects'; NQS: 'Non-Quantitative Subjects'; L&M: Law and Medicine (see section 4.2). Social Classes: I (professional), II (intermediate), IIINM (skilled non manual), IIIM (skilled manual), IV (partly skilled), V (unskilled). Standard errors and confidence intervals were computed using the delta method and Z critical values.

Table 5: Average gross weekly occupational earnings of leaving students by subject - 8 groups

Cohorts of leaving students	QS								NQS				L&M			
	Sciences		Maths		Technical		Economics and Business		Soft SS		Art & Humanities		Law		Medicine	
	mean	rank	mean	rank	mean	rank	mean	rank	mean	rank	mean	rank	mean	rank	mean	rank
1985	422	1	476	6	443	4	475	5	429	2	431	3	504	7	605	8
1986	458	1	519	5	482	4	528	6	465	2	469	3	553	7	650	8
1987	506	1	571	5	530	4	589	6	523	2	525	3	647	7	730	8
1988	565	2	641	6	601	4	635	5	563	1	569	3	704	7	832	8
1989	617	1	700	6	656	4	699	5	628	3	627	2	829	7	899	8
1990	647	1	722	6	678	4	701	5	656	2	663	3	988	8	977	7
1991	682	2	743	6	720	4	725	5	675	1	696	3	984	7	1082	8
1992	706	2	766	6	744	4	754	5	701	1	722	3	1019	8	1103	7
1993	717	1	794	5	775	4	796	6	729	2	754	3	1027	7	1122	8

Note. The table reports average gross weekly occupational earnings (from the New Earnings Survey) of leaving students by subject and the rank for each year. Average earnings are computed on the estimation samples (i.e. non-mature, non-overseas, unmarried students studying for a first degree qualification and with A-levels). For the definition of the subject groups see section 5.1.

Table 6: Predicted probabilities, with standard errors and 95% confidence intervals, of enrolling in the different study fields by social class - flexible-threshold ordered probit model, 1985 USR data

		Social class								
Subject		I	II	IIINM	IIIM	IV	V	Armed forces	I.D.	Non workers
	Prob	0.22	0.24	0.23	0.24	0.24	0.26	0.22	0.25	0.23
Sciences	left	0.21	0.22	0.22	0.22	0.22	0.22	0.19	0.23	0.20
	right	0.24	0.25	0.25	0.26	0.26	0.30	0.26	0.27	0.26
	Prob	0.09	0.10	0.10	0.10	0.10	0.10	0.12	0.10	0.11
Soft Social Sciences	left	0.08	0.08	0.08	0.08	0.09	0.07	0.09	0.08	0.09
	right	0.11	0.11	0.11	0.11	0.12	0.12	0.14	0.11	0.13
	Prob	0.25	0.26	0.25	0.25	0.26	0.25	0.26	0.25	0.26
Art and Humanities	left	0.24	0.24	0.24	0.23	0.24	0.22	0.23	0.23	0.24
	right	0.27	0.27	0.27	0.27	0.27	0.28	0.29	0.27	0.29
	Prob	0.12	0.11	0.12	0.12	0.12	0.10	0.14	0.12	0.11
Technical	left	0.11	0.10	0.10	0.11	0.10	0.07	0.11	0.10	0.09
	right	0.13	0.13	0.13	0.13	0.13	0.13	0.16	0.14	0.13
	Prob	0.08	0.09	0.10	0.09	0.09	0.09	0.08	0.09	0.08
Economics and Business	left	0.07	0.08	0.08	0.07	0.07	0.07	0.06	0.07	0.06
	right	0.10	0.10	0.11	0.10	0.10	0.12	0.10	0.11	0.10
	Prob	0.05	0.05	0.05	0.06	0.05	0.07	0.05	0.05	0.07
Maths	left	0.04	0.04	0.04	0.04	0.04	0.04	0.03	0.04	0.05
	right	0.06	0.06	0.06	0.07	0.06	0.09	0.07	0.06	0.09
	Prob	0.07	0.06	0.05	0.06	0.05	0.05	0.04	0.06	0.06
Law	left	0.05	0.04	0.04	0.04	0.03	0.02	0.02	0.04	0.03
	right	0.08	0.07	0.07	0.08	0.06	0.07	0.06	0.08	0.08
	Prob	0.11	0.09	0.10	0.09	0.10	0.09	0.09	0.09	0.09
Medicine	left	0.09	0.08	0.08	0.08	0.08	0.06	0.07	0.07	0.07
	right	0.12	0.11	0.11	0.11	0.12	0.11	0.11	0.10	0.11

Note. QS: 'Quantitative Subjects'; NQS: 'Non-Quantitative Subjects'; L&M: Law and Medicine (see section 4.2). Social Classes: I (professional), II (intermediate), IIINM (skilled non manual), IIIM (skilled manual), IV (partly skilled), V (unskilled). Standard errors and confidence intervals were computed using the delta method and Z critical values.

Table 7: Predicted probabilities, with standard errors and 95% confidence intervals, of enrolling in the broader subject groups by social class - flexible-threshold ordered probit model, 1985 USR data

Social Class	Subject group								
	QS			NQS			L&M		
	Prob	left	right	Prob	left	right	Prob	left	right
I	0.48	0.46	0.50	0.35	0.33	0.36	0.17	0.15	0.20
II	0.50	0.48	0.52	0.35	0.34	0.37	0.15	0.13	0.17
IIINM	0.50	0.47	0.52	0.35	0.33	0.37	0.15	0.13	0.18
IIIM	0.50	0.48	0.52	0.35	0.33	0.37	0.15	0.13	0.18
IV	0.49	0.47	0.52	0.36	0.34	0.38	0.15	0.12	0.17
V	0.52	0.47	0.57	0.35	0.31	0.38	0.13	0.09	0.18
Armed forces	0.49	0.45	0.53	0.38	0.35	0.41	0.13	0.09	0.17
Inadequately described	0.51	0.48	0.54	0.34	0.32	0.37	0.15	0.12	0.18
Non workers	0.49	0.45	0.52	0.37	0.34	0.40	0.14	0.11	0.18

Note. QS: 'Quantitative Subjects'; NQS: 'Non-Quantitative Subjects'; L&M: Law and Medicine (see section 4.2). Social Classes: I (professional), II (intermediate), IIINM (skilled non manual), IIIM (skilled manual), IV (partly skilled), V (unskilled). Standard errors and confidence intervals were computed using the delta method and Z critical values.