Fostering Educational Enrolment Through Subsidies: the Issue of Timing

Mario Fiorini *
University College London

November 2006

Abstract

In this paper we build a dynamic structural model of educational choices in which cognitive skills shape decisions. The model is estimated by maximum likelihood using cohort data where individuals are observed from birth onwards. These data are unique in that they include cognitive skills test scores collected as early as age 7. We then simulate the effect of two educational subsidies equal in cost but different in the timing of disbursement. The first consists of grants assigned directly to individuals aged between 16 and 18. The second is assigned to the parents earlier on, when the cohort is still in its childhood. The latter subsidy affects cognitive skills accumulation and in turn educational choices. Our results suggest that a direct grant in the form of a tuition subsidy might be more efficient even in the absence of short term financial constraints. Although cognitive skills accumulated during childhood play a key role in the educational decisions, an unconditional financial subsidy to parents is not the best policy. The results do not call a halt to investments in cognitive skill accumulation during childhood, but recommend that such investments should be well structured and ensure a high return.

Keywords: educational decisions, dynamic structural estimation, tuition subsidy, parental income subsidy.

JEL classification: I21, I28, J24, J31.

*I wish to thank Jerome Adda for his constant encouragement and support. I also thank Lars Nesheim, Richard Blundell, Pedro Carneiro, Giacomo De Giorgi, Marco Cozzi and Katrien Stevens for helpful comments.

1 Introduction

The objective of this paper is to investigate how educational choices are affected by cognitive skills accumulated during childhood. In particular, we address a specific policy question *"If the Government aims to increase enrolment in post-compulsory education which policy is more effective? should the Government subsidize parents at an early stage, to increase their investment in the child's cognitive skills during childhood, or should it instead subsidize individuals directly through grants in post-compulsory education? "*.

The relevance of this question is proved by recent policies introduced in the United States and in the United Kingdom. Specifically, in September 2004 the U.K. has introduced the Education Maintenance Allowance (EMA) to increase participation in education at the age of 16, the minimum school leaving age. In 2006 the University tuition fee system has changed, the so called Top-Up fees project, with a combination of higher tuition, loans and grants to induce enrolment in Higher Education.

Previous studies have tried to evaluate the effectiveness of financial incentives and short-term liquidity constraints at age 16 or later, but rarely found them to be relevant for educational and occupational choices. Their concluding remarks pointed to ex-ante heterogeneity as the main determinant of these choices. This ex-ante heterogeneity has always been treated as exogenous and called skill endowment. [1]

Our contribution is to test the effect of early policy intervention on cognitive skills and evaluate the impact on educational decisions. To tackle this issue we exploit the information provided by the National Child Development Survey (NCDS), a U.K. cohort study following individuals from birth onwards that collected information on skills and parental background information at different stages of childhood. We are therefore able to observe the skills accumulation process together with educational and occupational choices up to the age of 41. These data are unique, because they follow individuals from birth while the majority of other cohort studies, such as the National Longitudinal Survey of Youth (NLSY), follow individuals from the age of 14 or later, therefore lacking childhood information. The NCDS data give us the unique opportunity to test the effect of variation in cognitive skills measured as early as age 7. [2]

In our approach we consider both the investment of parents into the child's cognitive skills, through a production function capturing the causal effect of parental income on skills, and the educational choices determined by preferences and accumulated skills.

We model the educational choices with a dynamic structural model. The advantage of this approach lies in the possibility of inferring the preferences and beliefs of individuals, modelling the selection into schooling based on observables and unobservables and leaving us the option of evaluating the effect on educational decisions of changes in key parameters such as the distribution of cognitive skills and tuition fees.

Since we do not have a plausible exclusion restriction, identifying the return to parental income in cognitive skills is problematic. What we do instead is to use the structure of the model and the comparison of the two education subsidies, the tuition subsidy (grant) and the parental income subsidy, to identify which return to parental income would make the two subsidies equivalent. [3] We focus on men as the educational decisions and wage outcomes are gender specific. We further restrict the sample by excluding self-employed individuals since their wages are not always well reported.

[1] In this paper we distinguish between ability, that we consider innate, and skills, that are instead the result of ability and the learning process. Therefore ability and skills are identical only at birth.

[2] The NCDS follows individuals born in 1958 matching closely the NLSY79 which follows individuals born between 1957 and 1965.

[3] The Data include some measure of investment such as parental interest in the child's education (assessed by school teachers) but, the aim of the paper being to simulate alternative policies, it is not straightforward to imagine how this form of parental investment could be affected by Government policy. What we have in mind are financial forms of parental investments that could be changed through monetary subsidies.

Finally, our set up is a partial equilibrium model, and therefore does not take into account the consequences of increased skill levels or higher qualifications on their relative returns. As an implication our model can be used to assess the effect of only relatively small changes in the stock of human capital.

When we target an increase in Higher Education (college or equivalent) enrolment equal to 1% of the population, our results indicate that a subsidy (grant) at the age of 18 is the most efficient way. The same subsidy, given to parents when the individual is still in its childhood, would increase cognitive skills and in turn enrolment into Higher Education, but not as much as a direct Higher Education subsidy at the age of 18. The same conclusion is reached if we instead target an increase in education enrolment between age 16 and 18.

Our result does not imply that additional investment in cognitive skill accumulation is wrong, but that such investments should be well structured and ensure a high return. Otherwise they risk being inefficient.

The paper is organized as follows: In section 2 we review the main findings of the literature. The literature on the economics of education being quite vast, we focus on those papers investigating the importance of short term financial constraints and of comparative advantage, which are the most relevant to our policy question. Section 3 describes the UK education system. Section 4 presents the cohort data that we use. In Section 5 we go through the economics of the model, while in Section 6 and Section 7 we explain respectively its identification and estimation. Section 8 first presents the maximum likelihood estimates and the model fit, and then it describes our simulations. Section 9 concludes.

2 Related Literature

The recent US literature on educational choices has looked at the importance of financial incentives and budget constraints in schooling decisions.

Keane and Wolpin (1997) investigate the educational and occupational decisions of a US male cohort born in the 60's (NLSY 79) in a dynamic discrete choice structural model estimating the impact of a tuition fees subsidy on the college participation decision. In the model individuals are not financially constrained and from age 16 onwards decide year by year whether to stay in education, stay home, work in a white or blue collar occupation. The model allows for ex-ante (age 16) heterogeneity in endowment. Keane and Wolpin find that a $2000 college fees subsidy would increase high school graduation by 3.5 percentage points and increase college graduation rates by 8.4 percentage points. However, when analyzing the life-time utility effects, those who would benefit most from the subsidy are the individuals with high endowment for school education and white collar occupations, who would have gone to college even without the subsidy. Those induced to go to college by the subsidy are individuals with low endowment for education and a comparative advantage in blue collar occupations, with only a minor increase in their lifetime utility. This is due to part of the subsidy being spent to compensate the pre-policy larger utility from a no-college choice. They conclude that ex-ante heterogeneity in skill endowment is a major determinant of responses to subsidies and effects in lifetime utility.

Keane and Wolpin (2001) extend their previous work to account for financial constraints at the age of 18, with the model also allowing for parental transfers from age 16 onwards and marriage as important factors affecting the decisions. They find that borrowing constraints exist and are tight but have a limited effect on college attendance decisions, since individuals adjust their behavior working part-time or reducing consumption while at school. Hence, subsidizing poor parents would have little effect on college participation. The latter results support the hypothesis that ex-ante heterogeneity plays a central role in educational and occupational decisions.

Cameron and Heckman (1998) use an ordered choice dynamic model to explore whether the importance of family background factors in educational decisions decreases as individuals move towards higher grades. They look at five US cohorts born between 1907 and 1964. They find that these background factors have a rather constant importance across cohorts, but that family income is not so relevant once controlling for observed cognitive skills. Their conclusion is that parental factors are important mainly because they shape child's skills and taste for education early in life, and these latter characteristics determine educational choices.

Cameron and Heckman (2001) estimate a dynamic model of schooling attainment in the US, investigating the sources of educational disparities between Black, Hispanic and White Males. While it is often found that these disparities are linked to parental income differentials, the paper tests whether this effect is due to long-term effects or short term financial constraints. They estimate the model separately for the 3 ethnic groups and find that parental income is important, but its effect is largely diminished once they control for AFQT scores. They also test the effect of variation in costs, equalizing the tuition fees across ethnic groups, but again do not find large effects. On the other end, equalizing the AFQT, would lead to blacks and hispanics having higher enrolment rates than whites.

Carneiro and Heckman (2002) look at the same US cohort as the Keane and Wolpin studies, and investigate in a static model the importance of short run and long run factors influencing the college attendance decision. Short run factors are associated with liquidity constraints at age 18, while long run factors are linked to permanent differences due to parental background. Once again, conditioning on skills measured in early teenage years, short term constraints play only a minor role. Their results suggest that at most 8% of American youth face short term liquidity constraints that affect post secondary schooling.

Cameron and Taber (2004) measure the importance of borrowing constraints on education decisions. Their intuition is that opportunity costs and direct costs of schooling affect borrowing constrained and unconstrained persons differently. Direct costs need to be financed during school and impose a large burden on credit-constrained students. By contrast, gross forgone earnings do not have to be financed. They explore this idea using both a reduced form IV strategy and a structural model approach. However, in no case they find evidence of borrowing constraints.

Dearden, McGranahan, and Sianesi (2004) replicate the analysis of Carneiro and Heckman (2002) using UK data, the National Child Development Survey (NCDS) data and find that once controlling for test scores in mathematics and reading, individuals do not seem to suffer from short-term credit constraints. Their findings suggest that policies aimed at reducing the impact of credit constraints on education decisions should target individuals at the age of 16 (or possibly earlier) when staying-on decisions are made, rather than at age 18 when individuals are making Higher Education decisions. [4] However, given the reduced form approach and no information on schooling direct costs, they can not simulate the possible impact of a Government financial subsidy to parents or individuals on their educational decisions.

Attanasio, Fitzsimons, and Meghir (2004) apply a dynamic structural model with ex-ante (age 16) heterogenous individuals to estimate the impact of the Education Maintenance Allowance on the decision to either stay in education, stay in education and work part time, or leave education and work full time. They use the data provided by the experiment conducted in 1999 when the EMA was introduced in ten Local Education Authorities. This survey followed individuals for only 3 consecutive years from age 16. They find that the EMA program increased participation in education without a part time job from age 16 to 18 and increasing the generosity of the EMA would further augment such participation. However, the EMA would have only negligible effects on the participation in Higher Education. [5]

All together these studies provide a strong indication that policies aimed at increasing education attendance through monetary subsidies might not be very effective. Only a small fraction of individuals

[4] An example of this policy is the Education Maintenance Allowances programme.
[5] The EMA was only given until age 18.

appears not to enroll because of short term financial constrains, and those who enroll because of the subsidy do not have big gains in utility. Instead, these papers point to the comparative advantage hypothesis as originally described by Roy (1951) and subsequently by Willis and Rosen (1979). If an individual has accumulated enough skills then staying in education will be rewarding, while if the individual did not accumulate the skills staying on would actually lead to lower utility than otherwise obtained entering the labor market immediately.

An education subsidy could change the participation decision for those at the margin, but part of the subsidy would be lost in compensating the difference in utility caused by the comparative advantage.

The policy recommendation for a Government aiming to increase enrolment into education is to intervene not at age 16-18, when staying-on decisions are made, but during childhood, when individuals accumulate their skills, cognitive and non-cognitive. The same subsidy given at the age of say 11 could foster the accumulation of skills and be therefore more effective on schooling choices then if given at age 18. Nevertheless, these studies do not compare the effectiveness of the two alternative policies: the fees subsidy (grant) and alternative early intervention.

Estimating the impact of Government intervention so early in life is, however, a hard task given the shortage of surveys monitoring parental decisions and skills accumulation. The NLSY longitudinal data used in many of the papers mentioned above follow individuals at best from the age of 14 onwards, with a measure of skills given by the AFQT test score and with parental income measured at the age of 16. Therefore it does not allow testing the impact of subsidy to individuals or their parents anytime before.

3 The UK Education System

Before introducing the model, we give a brief overview of the UK Education System in order to understand the individual decision process and the assumptions behind our model.

We focus on the English and Welsh education system, which are identical, and omit the Scottish Education System, which is slightly different from the previous ones. This choice is driven by the difficulty of modelling the educational decisions for both types of systems.

Schooling is compulsory up to the age of 16, when individuals can, at the end of the scholastic year, stay in education or enter the labor market. If they stay, there are two main educational paths that they can follow: the Academic and the Vocational one. We describe both in turn although they are not necessarily mutually exclusive, i.e. individuals might take both Academic and Vocational Qualifications.

We discuss the system faced by individuals born in 1958 which, with some differences, also reflects the current one.

3.1 The Academic Path

The Academic path is mainly full-time. Those who stay on at age 16 enrol for the O Levels or CSE qualifications, which are taken immediately at the end of the scholastic year. These students are still aged 16 when they obtain the qualification. O Levels are single subject examinations reflecting the single disciplines of the university departments and faculties. They are designed for more able secondary school students and are necessary for progression into further education (A-level). The Certificate of Secondary Education (CSE) qualification was intended for students whose skills were not considered sufficient for O level courses. Nevertheless, there was an overlap between these two types of certificates in that a CSE grade 1 result was regarded as equivalent to an O-level pass. Even though there was no formal requirement, students would be expected to pass at least 5 O Levels graded A-C or CSE graded 1st, in order to stay in education afterwards.

In the Autumn term of the same year, those who successfully obtained 5 or more O Levels/CSE can enrol for A Levels. These would last 2 years, until individuals are aged 18. Advanced levels have their origins in qualifications constructed by groups of universities in the first half of the century, designed to identify candidates suitable for degree courses and to provide a foundation for advanced teaching in a single subject area to degree level. These characteristics still distinguish A levels. They are still awarded by independent examining bodies with close links to the universities. Passing 2 A-level constitutes the minimum level required for entry in Higher Education. Normally two or three A-levels are studied.

Once the student has completed A Level, he can gain admission to the Universities, Polytechnic or Colleges of Higher Education where a first degree is obtained. The time needed to gain a degree varies by subject but in the majority of cases it takes 3 years.

Therefore a student who completes the Academic path with no interruption will normally enter the labor market at the age of 21 or 22. [6] Hereafter we use OL, AL and HE for O Levels, A Levels and Higher Education respectively. [7]

3.2 The Vocational Path

The Vocational Path is quite different from the Academic one, mainly in the types of qualification awarded and in the timing. Vocational qualification are more specialistic and often linked to the acquisition of a competence requested for a specific kind of job. They range from advanced food technology, catering degrees to lower levels dog grooming and cake decoration ones. Some of these qualification are taken while in full-time education while others are taken later on in life, even after entering the labor market. Yet, vocational qualifications are grouped in O Level equivalents, A Level equivalents and Higher Education equivalents to match them with the Academic ones.

4 Data and descriptive statistics

The National Child Development Study (NCDS) targets over 17,000 babies born in Britain in the week 3-9 March 1958. Surviving members of this birth cohort have been surveyed on six further occasions in order to monitor their changing health, education, social and economic circumstances: in 1965 (age 7), 1969 (age 11), 1974 (age 16), 1981 (age 23), 1991 (age 33) and 1999 (age 41). At the age of 7, 11 and 16 mathematics, reading and general skills tests were taken by the cohort member. Information about parental background including education and occupation was also collected during those years, with a measure of parental income included at the age of 16. Moreover, in 1978, when individuals where aged 20, a survey was conducted among the secondary schools where they had taken their qualifications up to A Levels. This school survey allows to distinguish between enrolment in OL and AL and the actual achievement of the qualification.

Although the surveys were not conducted on yearly basis, information on labor market history, including employment status and occupation, was gathered for each month of the cohort members life from age 16 onwards through retrospective questions. Data on wages instead were only collected at the time of the interviews.

[6] Individuals might stay in education even further to complete a post-graduate degree. However this was very unlikely among the older cohorts such as the NCDS one.

[7] Another important feature of the British education system was the early selection of pupils into different types of schools. Prior to, and during the 1960s, the British education system was selective. Pupils were tracked into different schools, according to their ability at age 11. The most able, who passed an entrance examination at age 11, went to grammar schools. The rest went to secondary modern or technical schools. The selection of around only 20 per cent of the cohort for a grammar school education (which could lead to university entrance) was progressively challenged in the 1960s. From 1965 onwards (Circular 10/65), Labour governments encouraged local authorities to develop comprehensive schools which accepted all children from a neighborhood, regardless of ability. Comprehensive education slowly gained ground and by the end of the 1970s over 80 per cent of all children in maintained schools were in comprehensive schools.

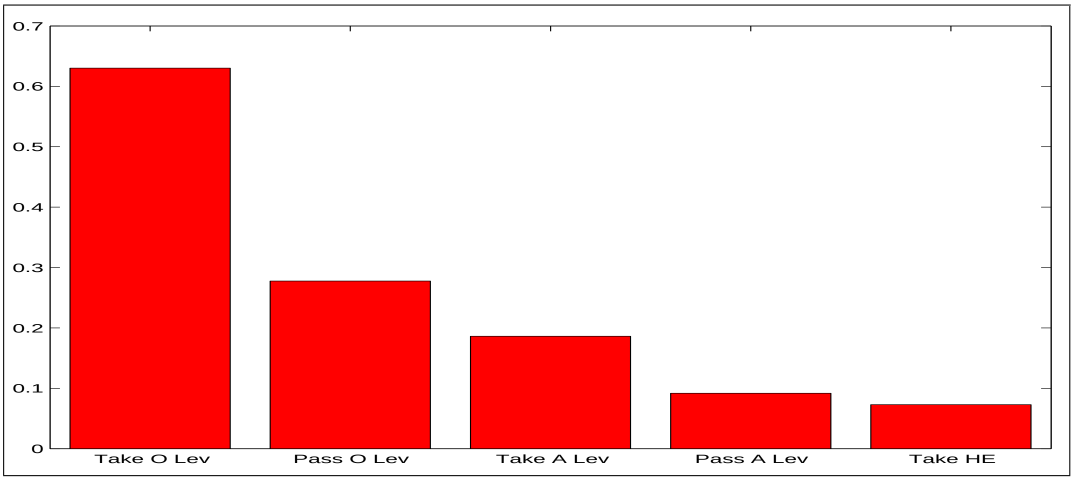

Figure 1: Educational choices

These data sets therefore bring together information on educational and occupational choices, skills and parental background measures collected in the childhood years of the cohort members.

We select all males for which we observe parental income, skills test scores at age 16 and educational choices.

Table 1 shows the coding and the composition of our sample by status and age. We select those individuals observed for at least two periods. Given the education system we start our analysis with the period April 1974 up to September 1974, entry 16^a in our table, when individuals were already 16 and could leave education. This is the moment we can observe their first decision. Because this period is mainly within the scholastic year and summer holidays, we assume that individuals either enrolled in education or stayed home, and are therefore assumed to be unemployed. The following period runs from October 1974 until September 1975, entry 16^b, when individuals could have studied A Levels, found a job or be unemployed. We classify individuals as employed if the number of months spent working was larger than time in unemployment. This period and all the following ones are yearly time periods, starting in October and ending in September of the following year. Above age 21, that is after O Levels, 2 years of A Levels and 3 years of Higher Education individuals are supposed to be in the labor market. [8] We do not model re-entry in education. Among those in full-time education more than 93% of the individuals did not have any break from full-time education, where a break is arbitrarily defined as a six months or longer period outside full-time education before going back. This percentage raises to 95% when we consider a break as 12 months or longer period. We drop those individuals who have a break of 12 months or longer.

From the data it is clear that the unemployment rate in the sample rose suddenly around the age of 22, from 5% to 10%, only to go down again two years later, at age 24. The work rate mirrored the unemployment one given that all individuals were already in education. This was the time when Margaret Thatcher was prime minister with improved productivity but soaring unemployment in the UK. What casts a shadow though is the peak in the series. Since individuals were interviewed at the age of 23 and 33, the reported unemployment from the age of 24 onwards comes from the age 33 survey, with a risk that individuals under-reported their unemployment in those years, particularly if that occurred for short periods. [9]

[8] By the age of 23, less then 1% reported having a post graduate qualification. This percentage raised to almost 4% by the age of 33. We exclude these individuals from our sample.

[9] We have done some investigation on this, though we could not find a reliable unemployment statistic in the early 80's

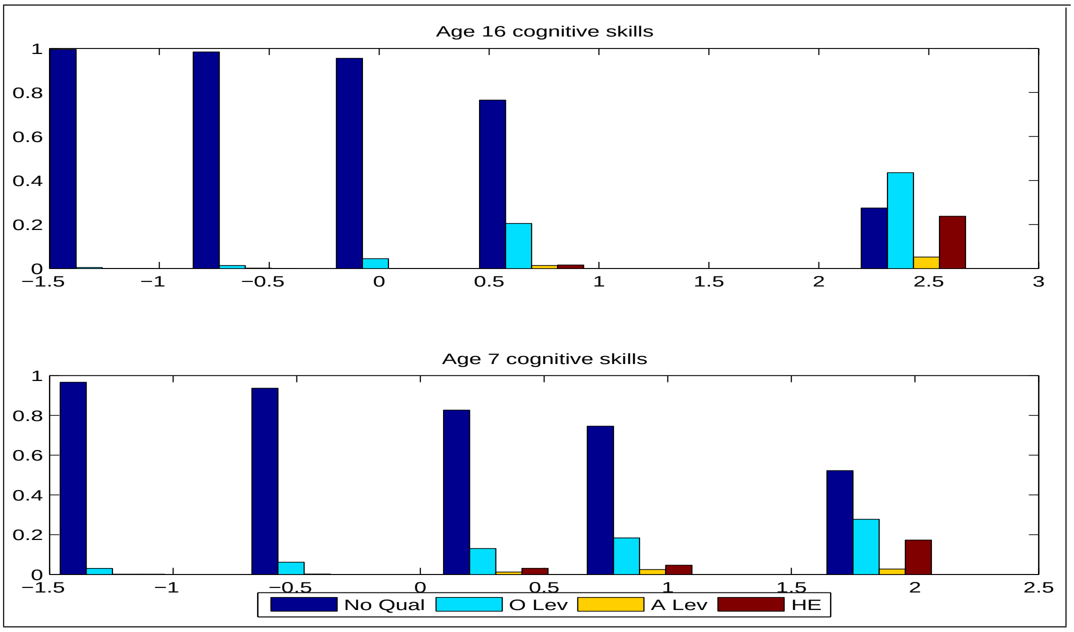

Figure 2: Qualifications by Skills.

Figure 1 shows the educational choices and the obtained qualifications. Although 63% of men in the sample enrolled for O Levels, only 28% obtained sufficient grades to progress to A Levels. Part of these individuals left education anyway. In the end only 7.3% of the sample obtained a HE qualification. Note that we do not distinguish between individuals who enrolled and individuals who successfully completed Higher Education. This is because, while data were collected from secondary schools, no data were collected from HE institutions. In our sample therefore we code as being enrolled all those individuals that reported having a HE qualification. [10]

Table 3 shows sample statistics for the mathematics and reading test scores at the age of 16 and 7, and for parental income at the age of 16. At age 7 individuals were administered four tests (reading, mathematics, copying and drawing) while only reading and mathematics tests were administered at age 16. Between age 7 and 16 the tests were changed to take into account the age difference. We only use the mathematics and reading scores. [11] In order to reduce the number of state variables we summarize the mathematics and reading test scores using the first principal component of the standardized test scores. The assumption that one factor captures cognitive skills is quite widespread in the literature and in line with the g theory used by Herrnstein and Murray (1994). Not reported here, the first principal component for the age 16 scores explains 82% of the total variance while the age 7 principal component explains 77 % of the total variance. The loading factors of the standardizes scores are identical and equal to 0.70711, both at age 16 and 7. Higher values of the principal component correspond to higher

for those born in 1958. (i.e. a cohort specific statistic.) Nevertheless using available statistics or some other data set such as the FES, it seems that my data underestimate unemployment by around 3% between the age of 24 to 26.

[10]We have also tried to investigate this issue using some self-reported data at age 23, but the results are not fully convincing. It seemed that up to 16% of individuals might have failed an Higher Education course. It is hard to say whether we could do more on this because not all individuals in the sample actually took part in the age 23 survey, and for the missing ones we would not know whether they failed an HE course.

[11]The questionnaires, including the tests can be downloaded from the UK data archive website. The age 7 questionnaire can be found at http://www.data-archive.ac.uk/doc/3148/mrdoc/pdf/a3148uab.pdf, while the age 16 questionnaire can be found at http://www.data-archive.ac.uk/doc/3148/mrdoc/pdf/a3148ucb.pdf.

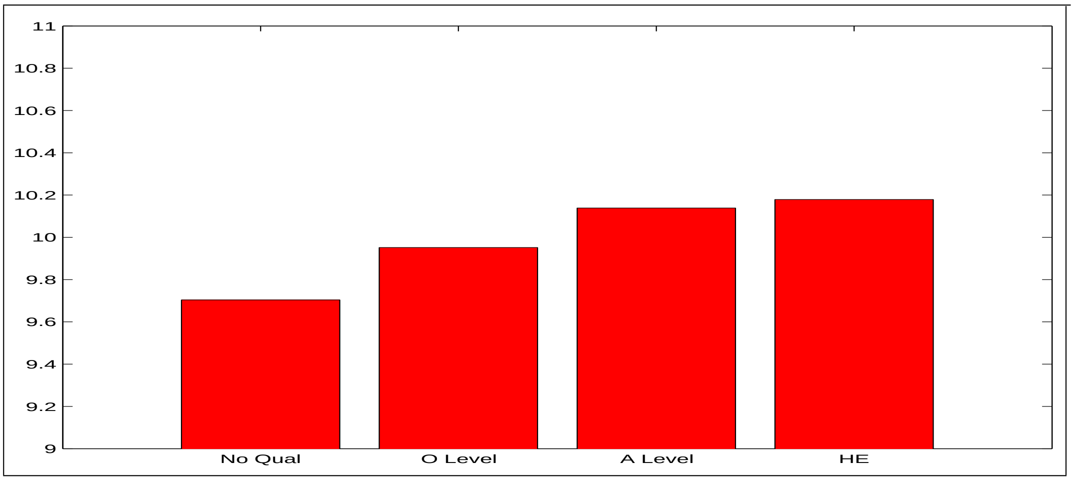

Figure 3: Log wages by qualification

scores. [12]

Figure 2 shows the obtained qualifications by principal component (quintiles). It is clear that both age 16 and age 7 cognitive skills are good predictors of educational achievement. Those in the lower quintiles have virtually no academic qualification.

Figure 3 reports mean log yearly wages by obtained qualification (in January 2001 prices). The wages are monotonically increasing. In table 2 we report separate statistics for the three age points when we can actually measure wages. In the first line we report the mean wage by age; as expected it is increasing as the individuals grow older. The following lines decompose the first one by highest qualification obtained. Higher qualifications are usually associated with higher wages.

The British NCDS cohort is comparable to the US National Longitudinal Survey of Youth (NLSY79) cohort, composed of individuals born between 1957 and 1965. The NCDS though followed the individuals from birth, while the NLSY first surveyed individuals in 1979 when they were between 14 and 22 years old. The NCDS study reports reading and mathematics test scores at the age of 7, 11 and 16. This is unique. The NLSY79 reports the AFQT test but this was administered in 1980 to all. Individuals in the NLSY were therefore aged between 15 and 23 when they took the test. The NCDS has also age 0, 7 and 11 information on parental background characteristics which is not present in the NLSY. [13]

5 The Model

In modelling the educational choices we only look at the Academic Path. This is because the qualifications are quite homogenous and the timing is very similar across individuals. In the data only 2% of individuals had Vocational qualifications without any Academic one at the age of 23. Conlon (2001) found that the wage returns to Academic qualifications are usually much larger than to Vocational qualifications. We choose to model an optimal stopping problem, where at the age of 16 (OL and AL) and

[12] The loading coefficient of the original, non-standardized scores were also very close and equal to 0.71 and 0.69 respectively for the math and reading scores.

[13] The Armed Forces Qualifications Test score (AFQT) is a composite score derived from select sections of the Armed Services Vocational Aptitude Battery (ASVAB), a battery of 10 tests that measure knowledge and skill in the following areas: (1) general science, (2) arithmetic reasoning, (3) word knowledge, (4) paragraph comprehension, (5) numerical operations, (6) coding speed, (7) auto and shop information, (8) mathematics knowledge, (9) mechanical comprehension and (10) electronics information.

18 (HE) individuals decide between staying on or leaving education and enter the labor market. The labor market is an absorbing state. Given the low number of individuals that re-enter education after a year long break, an optimal stopping model should match fairly well the choices made in the Academic path.

5.1 The Decision Process

At the age of 16 individuals are free to leave education. Before they take the decision they are assumed to know the probability of receiving a job offer (δ). They remain in education whenever the expected lifetime utility from schooling is higher than from entering the labor market. If they decide to stay in education they enrol in OL exams, with a certain probability (λ) they succeed and in the next period choose between AL or leaving education. If they fail ($1 - \lambda$) then they have to leave. As long as in education, this decision process is repeated up to HE, the highest possible level.

Once this is obtained individuals have to enter the labor market. Given that we focus on men, we assume that once out of education individuals always choose to work. Nevertheless, they could be unemployed if they do not receive a job offer (provided they were in education or unemployed in period $t-1$) or if they get fired (provided they were working in period $t-1$) which occurs with a probability (ϕ). The only decision in the model is between staying or leaving education. The work/unemployement status is not a decision.

This process being sequential, we model it as a discrete Markov decision process (DMDP) where at each point the choice depends only on the current level of the individual characteristics or state space. See Eckstein and Wolpin (1989) and Rust (1994) for a review of DMDP models and estimation strategies.

5.2 Instantaneous Utility

We follow Keane and Wolpin (1997) and model the rewards in monetary terms. This choice is dictated by the absence of variation in tuition fees in the data. [14] The linear utility in income implies that the individuals are not liquidity constrained and are risk neutral. In the conclusions we discuss possible implications of these assumptions. Individual and time specific subscripts are suppressed for clarity of exposition.

The reward from work is given by the annual wage:

$$R_W = w = exp\left\{\sum_{j=1}^{3}\alpha_{qj}Q_j + \alpha_x X + \sum_{j=1}^{3}\alpha_{xj}XQ_j + \alpha_{k2}K_{16} + \alpha_{k1}K_7 + \alpha_z + u\right\} \tag{1}$$

The effect of schooling is represented by the 3 dummies (Q_j), corresponding to obtained OL, AL and HE qualifications and their interaction with experience (X). The equation also includes skills at age 16 (K_{16}), age 7 (K_7). We interact qualification with experience to better fit the wage profile. We include K_7 since it is a good proxy for innate ability and in the policy counterfactual we test the effect of a change in its distribution.

The unobserved component includes a type specific constant α_z and measurement error u: $u \sim N(0, \sigma_u)$. We model the unobserved heterogeneity with a mixture model and Z types. [15] Provided

[14] Ichimura and Taber (2002) discuss semiparametric identification of tuition subsidy effects whenever the researcher can observe variation in tuition. While in England and Wales education at every stage was free, individuals enrolled in Higher Education were entitled to a grant linked to parental income. The amount of the grant was between 200 and 5000 pounds in 2001 prices. Since a tuition subsidy is equivalent to a grant, we could try to use variation across individuals. However, given its link to parental income, it is not easy to identify the effect of the grant net of income.

[15] See Everitt and Hand (1981), McLachlan and Peel (2000) and Heckman and Singer (1984) for a discussion of mixture models.

$\{\alpha_{k2}, \alpha_{k1}\}$ are positive and the return to schooling are monotonically increasing in Q, the exponential functional form ensures a positive cross-derivative between schooling and cognitive skills.

The reward from unemployment is

$$R_U = \overline{R} \tag{2}$$

that is individuals receive a fixed amount in unemployment benefit if they are unemployed. $\overline{R}$ is set equal to 3000 pounds, approximately the 1975 benefit in 2001 prices. Until 1995 the benefit was not a function of income. [16]

The latent reward from schooling is given by

$$Rs_j = \gamma_j + \gamma_z + \epsilon_j \quad j = \text{OL, AL, H.E.} \tag{3}$$

which indicates that the reward is qualification specific. In the case of O Levels, the first schooling stage, γ_{OL} is set equal to zero because we already have the type specific constant term. The reward is expressed in monetary terms and includes the costs as well as the consumption value of education as measured by the constant term γ_{j0} and γ_z, the persistent unobserved heterogeneity common to all schooling levels. The taste shock (ϵ) is assumed to be i.i.d. and normally distributed $N(0, \sigma_\epsilon)$. Tuition fees are imposed on the model but set equal to zero to reflect free education in the UK at that time. Note that there is no observed heterogeneity in the latent reward from schooling. Cognitive skills are excluded to help identification as we explain later on. This restriction holds if parental transfers and effort are not skill specific, or if they offset each other, i.e. less skilled individuals make more effort but also receive more resources from parents to compensate for their skill gap. We do not include parental income either. Carneiro and Heckman (2002) and Dearden, McGranahan, and Sianesi (2004) have found that parental income does not affect educational decsions once cognitive skills are controlled for. Moreover, it is a continuous variable and would increase computational time. [17]

Finally define the latent index generating the probability of obtaining a qualification once enrolled λ, the probability of receiving a job offer δ, the probability of being fired ϕ as

$$\begin{aligned}
\lambda_j^* &= \lambda_j + \lambda_{k2}K_{16} + \lambda_{k1}K_7 + \lambda_z + \varrho_\lambda \\
\delta^* &= \sum_{j=1}^{3} \delta_{qj}Q_j + \delta_{k2}K_{16} + \delta_{k1}K_7 + \delta_x X + \delta_z + \varrho_\delta \\
\phi^* &= \sum_{j=1}^{3} \phi_{qj}Q_j + \phi_{k2}K_{16} + \phi_{k1}K_7 + \phi_x X + \phi_z + \varrho_\phi
\end{aligned} \tag{4}$$

such that the unobserved heterogeneity enters each probability and ϱ is normal.

5.3 Solving the Sequential Decision Problem

Define the state space $\Omega = [K_{16}, K_7, Q, X, \epsilon, \alpha_z, \gamma_z, \lambda_z, \delta_z, \phi_z]$ where $[K_{16}, K_7, \epsilon, \alpha_z, \gamma_z, \lambda_z, \delta_z, \phi_z]$ are exogenous states in the model. Define the decision space $D = [Schooling, Non-Schooling]$, the status space $Z = [S, W, U]$ which differs from D since work and unemployment are not a choice. Finally define the choice set $C = \{|D(\Omega)|\}$ because individuals have a choice only if they are still in education and hold a qualification lower than HE. Individuals are therefore ex-ante heterogeneous in observed skills K,

[16] Source: Institute for Fiscal Studies , http://www.ifs.org.uk/ff/indexben.php

[17] Even though cognitive skills do not enter directly the utility of schooling they do affect schooling decisions as we explain later.

persistent unobserved heterogeneity $\{\alpha_z, \gamma_z, \lambda_z, \delta_z, \phi_z\}$ and serially uncorrelated unobserved taste shocks for education ϵ.

We characterize the problem of a finitely lived individual as maximizing the expected present value of lifetime rewards. The value function V of an individual of age a is defined as the solution to Bellman's equation

$$V(\Omega, a) = \max_{d \in C(\Omega)} \left[R(\Omega, d) + \beta \int_{\epsilon'} V(\Omega', a') \right] \tag{5}$$

where we integrate over the support of ϵ only since all the other states have deterministic dynamics. Equation (5) also highlights the advantage of not modelling the work/no-work choice since this way we only have a univariate integral.

5.3.1 Status Specific Value Functions

Define by V_S the value function from schooling, V_W the value function from work and V_U the value function from being unemployed. Then

$$V_{WU}(\Omega) = \delta(\Omega) V_W(\Omega) + [1 - \delta(\Omega)] V_U(\Omega) \tag{6}$$

is the value function if the individual exits education and enters the labor market.

$$V_{SWU}(\Omega) = \int_{\epsilon} max \{V_S(\Omega), V_{WU}(\Omega)\} \tag{7}$$

is the value function for an individual deciding between staying-on or leaving education.

Having in mind that OL, AL and HE have a different time length and that ϵ is updated every year, the value function from Schooling is therefore defined as:

$$\begin{aligned} V_{OL}(\Omega) &= R_{OL}(\Omega) + \beta \left\{ \lambda(\Omega) V_{SWU}(\Omega') + [1 - \lambda(\Omega)] V_{WU}(\Omega) \right\} \\ V_{AL}(\Omega) &= R_{AL}(\Omega) + \beta E_\epsilon[R_{AL}(\Omega)] + \beta^2 \left\{ \lambda(\Omega) V_{SWU}(\Omega') + [1 - \lambda(\Omega)] V_{WU}(\Omega) \right\} \\ V_{HE}(\Omega) &= R_{HE}(\Omega) + \beta E_\epsilon[R_{HE}(\Omega)] + \beta^2 E_\epsilon[R_{HE}(\Omega)] + \beta^3 \left\{ \lambda(\Omega) V_{WU}(\Omega') + [1 - \lambda(\Omega)] V_{WU}(\Omega) \right\} \end{aligned} \tag{8}$$

where for clarity we use Ω rather than Ω' if ϵ is the only state to be updated. An individual's value function is specified as the current reward plus future expected utility, which depends on the probability of success in school λ and the probability of receiving a job offer δ. The next period choice between school and labor market (V_{SWU}) vanishes when entering HE.

From our value function specification it can be seen that we update educational qualification to a higher level (Ω') only if the individual was successful once enrolled. Given what we said in section 3, this occurs if he obtained 5 or more OL-CSE at the first stage, 3 or more AL at the second stage. The implication is that enrolling for a qualification but achieving anything less than that has no effects on utility.

It might seem incorrect to assign an annual reward for O Levels even though they do not require an additional year of schooling. Nevertheless, we believe that the decision to take O Levels is actually taken around age 15 because it involves specific preparation to the exams. Hence, we still consider R_S as an annual reward, where the year goes from age 15 to 16. We instead set to zero the working reward for not taking O Levels, since the individual would formally still be in education or unemployed for a few months in the summer.

Value function from work:

$$V_W(\Omega) = R_W(\Omega) + \beta\left\{\phi(\Omega')V_U(\Omega') + [1-\phi(\Omega')]V_W(\Omega')\right\} \tag{9}$$

so an individual working has current utility given by the annual wage and future utility from work or unemployment, weighted by ϕ, the probability of being fired. Here Ω' captures the new level of experience.

Value function from unemployment:

$$V_U(\Omega) = R_U(\Omega) + \beta\left\{\delta(\Omega)V_W(\Omega) + [1-\delta(\Omega)]V_U(\Omega)\right\} \tag{10}$$

We do not assume any depreciation of experience so the state space does not change after a period of unemployment.

Note that there is no randomness in either V_W or V_U since both ϕ and δ do not contain any stochastic component and under the assumption of rational expectations the probabilities are always known by the individual.

5.4 Economics of the model

Being in school has a direct utility given by R_S, a cost given by the foregone earnings and a return given by the higher wage productivity. Taking O Levels and A Levels is also valuable because it allows access to higher qualifications (option value). Finally, schooling affects the probabilities of being employed in the labor market. Individuals select themselves in education based on their characteristics. Given the exponential form of R_W and conditional on $\nabla_K R_W > 0$, $\nabla_Q R_W > 0$ as it usually found in the literature, $\nabla_{QK} R_W > 0$. Individuals with high values of K have larger foregone earnings but also larger returns to schooling in the future. They will enroll in education as long as these returns are large enough.

The sign of λ, δ and ϕ gradients with respect to Q and K are also very important in driving the educational decisions. A large value of λ induces individuals to stay longer in education. Enroling in education but failing to get the qualification is very costly in the model because there is no change in human capital. The cost is the foregone wage weighted with the probability of finding a job. Thus if $\nabla_K \lambda > 0$ skilled individuals would be more likely to enrol. The effect of δ and ϕ depends on their interaction with the qualifications. If $\nabla_Q \delta > 0$ and $\nabla_Q \phi < 0$ then individuals have incentives to enrol. The sign of the cross-derivatives $\nabla_{QK}\delta$, $\nabla_{QK}\phi$ determines the selection based on observable skills.

Selection on unobserved heterogeneity works in a very similar way. High wage types (α_z) have larger forgone earnings but also larger returns to schooling if $\nabla_Q R_W > 0$. Enrolment in education is also caused by larger utility of schooling (γ_z), success at the exams (λ_z), larger probability of finding a job (δ_z) if $\nabla_Q \delta > 0$ and a lower probability of being fired (ϕ_z) if $\nabla_Q \phi < 0$.

This class of problems can easily be solved by backward induction. Our problem is particularly simple because u is just a measurement error and because decisions are made only at 3 points in life. Therefore the computation of the max in equation (5) occurs only at these points.

6 Identification

Magnac and Thesmar (2002) show that without unobserved heterogeneity these models are not non-parametrically identified as long as the following structural parameters are not set: the distribution of unobserved shocks, the discount rate, and the current and future preferences in one reference alternative. When unobserved heterogeneity is introduced, non-parametric identification is prohibitive unless very strong restrictions are imposed. We fix the discount factor β to 0.95 and we impose an exponential function for the utility of work in line with Keane and Wolpin (1997). In the utility of schooling we only impose the additivity of the shocks. Normality is imposed on measurement error, the schooling shock and over the transition probabilities λ, δ and ϕ. We make no assumption on the distribution of unobserved heterogeneity. Overall the model is heavily parametric because we are interested in the effect of different schooling levels and in the selection of individuals based on observable skills.

Parameters in R_W, ($\boldsymbol{\alpha}$), are identified from data on wages and the state variables skills, education and experience. The unobserved heterogeneity (α_z) is identified by cross-section variation in wages, conditional on the states, at each of the three wage points available in the data. Because we model selection on unobservables, the return to schooling, to observed skills and to experience are different from OLS ones. No parameter is estimated in R_U. In R_S the parameters ($\boldsymbol{\gamma}$) are estimated to match the proportions enrolling for each qualification. The $\boldsymbol{\sigma_\epsilon}$ vector is identified by a model's constraint: the net opportunity cost has a negative effect on the probability of enrolling. [18]

Identifying the scale of the parameters is necessary to test the effect of variation in tuition fees, given that we do not observe this variable nor the utility of schooling. The identification of $\boldsymbol{\sigma_\epsilon}$ is helped by excluding (K_{16}, K_7) from the utility of schooling. Formally this exclusion restriction is not needed given the non-linearity of V_{WU}, but without a restriction multi-collinearity would lead to large standard errors. The unobserved heterogeneity in the utility of schooling is identified by cross-section variation in schooling choices, conditional on the state variables, at each of the three schooling stages.

The parameters in λ, vector ($\boldsymbol{\lambda}$), are identified merging individual characteristics to the school data, the latter providing information on the number of OL and AL taken and those obtained. Here the unobserved heterogeneity is identified through cross-section variation, conditional on the state variables, in O Levels and A Levels exam success. The parameters in δ, vector ($\boldsymbol{\delta}$), and ϕ, vector ($\boldsymbol{\phi}$), are identified from yearly data on employment status. Unobserved heterogeneity is identified by cross-section variation in the transitions, conditional on the states, at each point in the labor market. Once again the returns to schooling, to skills and to experience take into account selection on unobservables and are therefore different from reduced form estimates.

7 Estimation

Define with L_w the likelihood from the wage density, with $L_\lambda, L_\delta, L_\phi$ the likelihoods from the probability of success in school, having a job offer and being fired, and with L_{OL}, L_{AL}, L_{HE} the likelihoods from the probability of enrolling in O Level, A Level and Higher Education. In the absence of unobserved heterogeneity and with uncorrelated error terms the log likelihood would be:

[18] To see clearly why, let us focus on the HE participation decision and assume for simplicity that the probability of success in school $\lambda = 1$ and the probability of having a job offer $\delta = 1$. Therefore an individual enrols in HE if:

$$R_{HE}(\Omega) + E\left[\sum_{t=1}^{2} \beta^t R_{HE}(\Omega)\right] + V_W(Q = HE, \Omega, a = 21) > R_W(Q = AL, \Omega) + \beta V_W(Q = AL, \Omega', a = 19)$$

Define the net opportunity cost as

$$A(\Omega) = R_W(Q = AL, \Omega) + \beta V_W(Q = AL, \Omega', a = 19) - V_W(Q = HE, \Omega, a = 21)$$

An individual with A Levels enrols in Higher Education with probability

$$1 - \Phi\left(\frac{A(\Omega) - \widetilde{R_{HE}}(\Omega) - E\left[\sum_{t=1}^{2} \beta^t R_{HE}(\Omega)\right]}{\sigma_{\epsilon,HE}}\right) \tag{11}$$

where $\widetilde{R_{HE}}$ is the utility of schooling net of the current realization of the shock. Since $A(\Omega)$ is determined by wage data and its coefficient is implicitly normalized to 1 in (11) by the model structure, $\sigma_{\epsilon,HE}$ is identified. The same reasoning can be generalized to identify σ_ϵ for the other educational categories.

$$
\begin{aligned}
\ell(\mathbf{\Omega_0}, \mathbf{\Theta}) &= \ln \prod_{i=1}^{N} (L_w \times L_\lambda \times L_\delta \times L_\phi \times L_{OL} \times L_{AL} \times L_{HE}) \\
&= \sum_{i=1}^{N} (\ell_w + \ell_\lambda + \ell_\delta + \ell_\phi + \ell_{OL} + \ell_{AL} + \ell_{HE})
\end{aligned} \tag{12}
$$

Given the additivity of $\ell(\mathbf{\Omega_0}, \mathbf{\Theta})$, estimation could be easily carried out by fast sequential maximum likelihood estimation using the backward induction nature of the problem. The $\{\boldsymbol{\alpha}, \boldsymbol{\lambda}, \boldsymbol{\delta}, \boldsymbol{\phi}\}$ parameter vectors could be identified by running separate likelihood maximization of L_w, L_λ, L_δ and L_ϕ. Then we could sequentially maximize $L_{HE}(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\delta}}, \hat{\boldsymbol{\phi}}, \boldsymbol{\gamma_{HE}})$, solving the problem of an individual with A Levels who chooses whether to enroll in Higher Education, given the known wage returns to schooling and experience, and job probabilities. With $\boldsymbol{\gamma_{HE}}$ estimated, $L_{AL}(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\delta}}, \hat{\boldsymbol{\phi}}, \boldsymbol{\gamma_{\hat{HE}}}, \boldsymbol{\gamma_{AL}})$ and then $L_{OL}(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\delta}}, \hat{\boldsymbol{\phi}}, \boldsymbol{\gamma_{\hat{HE}}}, \boldsymbol{\gamma_{\hat{AL}}}, \boldsymbol{\gamma_{OL}})$ could be maximized sequentially with the same logic. The inconsistent standard errors, due to the estimation error, could be corrected with one Newton step over the whole likelihood. (see Rust (1994) for a discussion on sequential estimation.)

However, when we introduce unobserved heterogeneity the log likelihood becomes:

$$
\begin{aligned}
\ell(\mathbf{\Omega_0}, \mathbf{\Theta}, \boldsymbol{\pi}) &= \ln \prod_{i=1}^{N} \left(\sum_{z=1}^{Z} \pi_z (L_w^z \times L_\lambda^z \times L_\delta^z \times L_\phi^z \times L_{OL}^z \times L_{AL}^z \times L_{HE}^z) \right) \\
&= \sum_{i=1}^{N} \ln \left(\sum_{z=1}^{Z} \pi_z (L_w^z \times L_\lambda^z \times L_\delta^z \times L_\phi^z \times L_{OL}^z \times L_{AL}^z \times L_{HE}^z) \right)
\end{aligned} \tag{13}
$$

with π_z being the proportion of individuals of type z. Equation (13) can no longer be estimated sequentially, because now we have the log of a sum.

However, Arcidiacono and Jones (2003) show how to extend the Expectation Maximization (EM) algorithm to solve the likelihood maximization sequentially. The EM algorithm consists of two steps:

- E Step. Compute the conditional probability of being the zth type:

$$
P(z|\mathbf{\Omega_0}, \mathbf{\Theta}, \boldsymbol{\pi}) = \frac{\pi_z (L_w^z \times L_\lambda^z \times L_\delta^z \times L_\phi^z \times L_{OL}^z \times L_{AL}^z \times L_{HE}^z)}{\left[\sum_{z=1}^{Z} \pi_z (L_w^z \times L_\lambda^z \times L_\delta^z \times L_\phi^z \times L_{OL}^z \times L_{AL}^z \times L_{HE}^z)\right]} \tag{14}
$$

- M Step. Estimate $\mathbf{\Theta}$ by maximizing the expected likelihood function and update the vector of unconditional probabilities $\boldsymbol{\pi}$, holding the conditional probabilities fixed.

$$
\begin{aligned}
\mathbf{\Theta} &= \arg\max \sum_{i=1}^{N} \sum_{z=1}^{Z} P(z|\mathbf{\Omega_0}, \mathbf{\Theta}, \boldsymbol{\pi}) \ln \left(L_w^z \times L_\lambda^z \times L_\delta^z \times L_\phi^z \times L_{OL}^z \times L_{AL}^z \times L_{HE}^z \right) \\
&= \arg\max \sum_{i=1}^{N} \sum_{z=1}^{Z} P(z|\mathbf{\Omega_0}, \mathbf{\Theta}, \boldsymbol{\pi}) \left(\ell_w^z + \ell_\lambda^z + \ell_\delta^z + \ell_\phi^z + \ell_{OL}^z + \ell_{AL}^z + \ell_{HE}^z \right)
\end{aligned} \tag{15}
$$

$$
\pi_z = \frac{\sum_{i=1}^{N} P(z|\mathbf{\Omega_0}, \mathbf{\Theta}, \boldsymbol{\pi})}{N} \tag{16}
$$

The two steps are repeated until convergence is achieved. The EM algorithm restores the additive separability at the maximization step in equation (15). Therefore estimation can be done sequentially. Arcidiacono and Jones (2003) name this procedure Expectation Sequential Maximization because it applies the EM algorithm to a sequential maximization problem. They show that this method produces consistent estimates of the parameters with large computational savings.

8 Results

8.1 The Dynamic Model

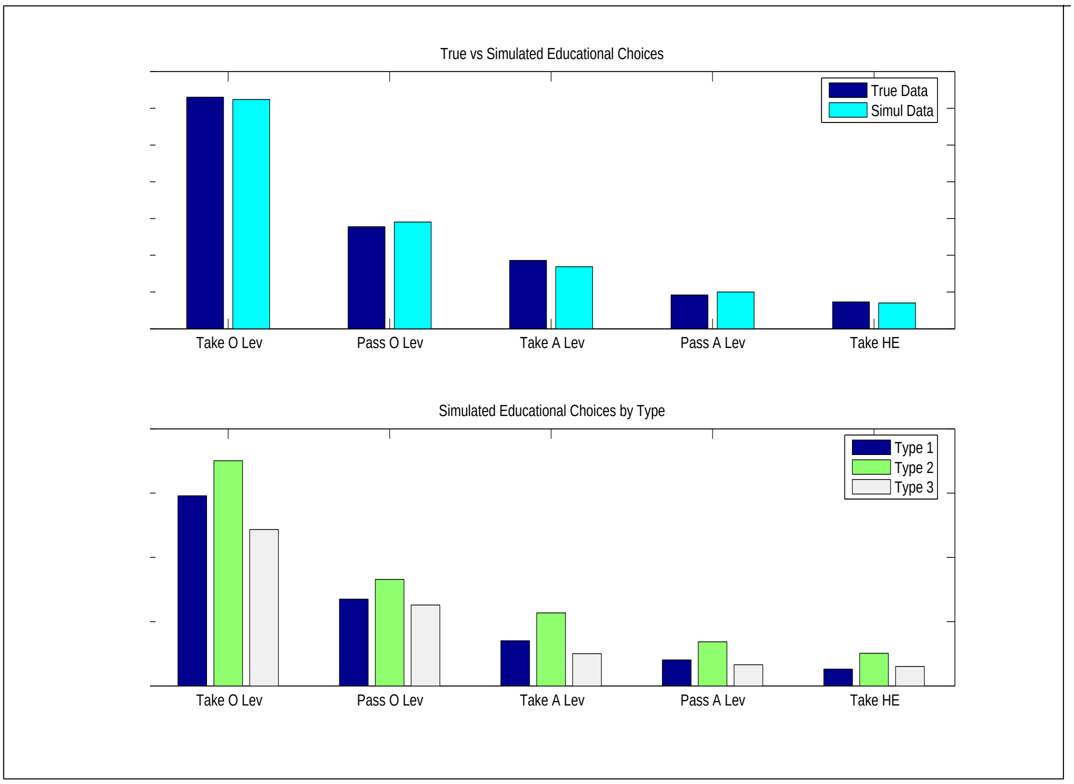

Figure 4: True and simulated educational choices

We first show the fit of the model and then present the estimated coefficients. A likelihood ratio test suggests that 3 types are sufficient to capture the unobserved heterogeneity: 60% of the individuals in our cohort are estimated to be type 1, 35% of type 2 and 5% of type 3. So far we have remained silent on what is unobserved heterogeneity. In structural schooling models like ours, researchers call it skill endowment (Keane and Wolpin (1997), Keane and Wolpin (2001)), or ability (Belzil and Hansen (2002) and Arcidiacono (2005)). There is also growing attention to the importance of non-cognitive skills as determinants of labor market outcomes and schooling (see Heckman, Stixrud, and Urzua (2006)). Since in our model we include early measures of cognitive skills among the observed characteristics, we do not consider the unobserved heterogeneity as ability. Rather we prefer to think of unobserved heterogeneity as those non-cognitive skills that are uncorrelated with cognitive ones. Figure 4 shows the true and simulated educational choices. The model fits quite well the educational choices. From the figure we can see that type 2 are the more likely to go into education, followed by type 1. Table 4 reports the fit by education, work and unemployment status. The model fits less well the way individuals

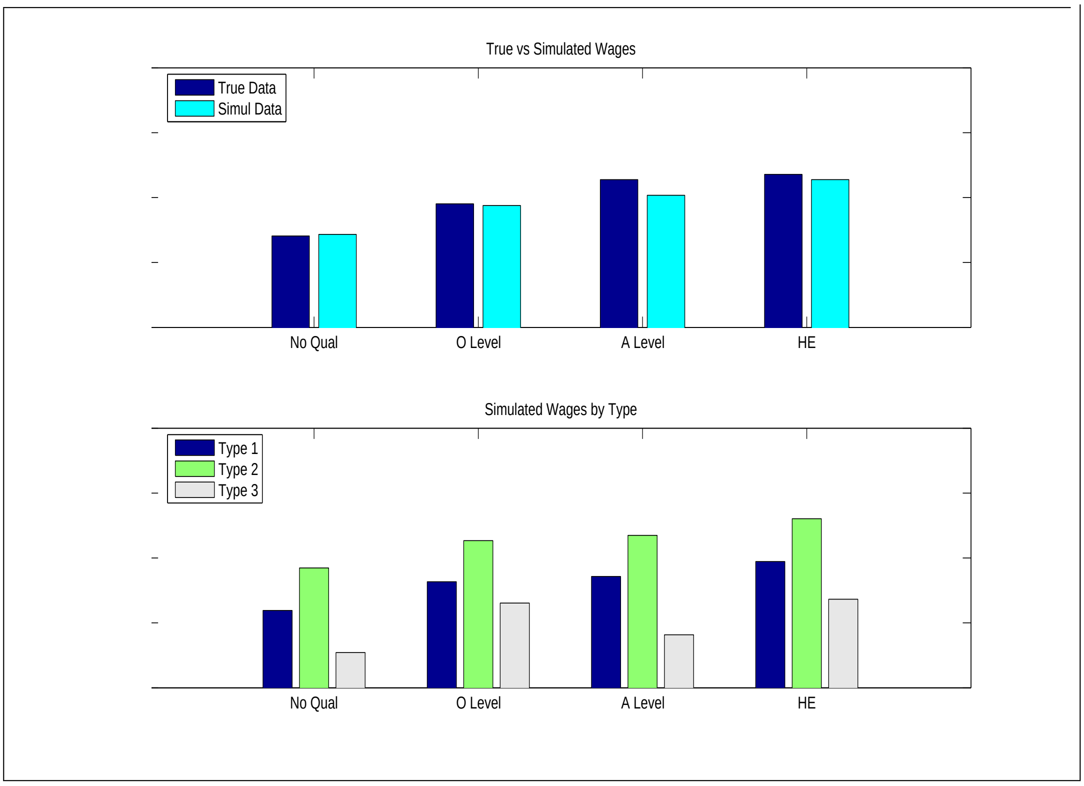

Figure 5: True and simulated log wages by qualification

allocate themselves between work and unemployment. The reason for this latter result is the peak in the unemployment series in period 8, at the age of 22, which the model can not really fit given that the job offer (δ) and job firing (ϕ) probabilities are not period specific.

However we only model the staying-on choice. Having left the education system, work or unemployment are not choices. Moreover, our main focus is to predict the decision of leaving school, which our model does reasonably well excluding the case of A Levels enrolment where there is a 1.7% gap between true and predicted data.

Figure 5 shows the wage fit. The model predicts quite well the wages by education groups with the exception of the A Level case, where there is a 2800 pounds difference. Unfortunately, given the low enrolment rates in A Level and Higher Education, we do not have many wage observations for these qualification so the imperfect fit is not fully surprising. Table 5 shows the true and predicted wage by time and qualification. The average wage is quite well predicted for all the 3 age points available in the data. However, when we decompose it by qualification, our model slightly under-estimates the unconditional wage return to A Levels and HE at the age of 33 and 41.

Figure 6 and 7 show how the model fits the educational outcomes by cognitive skill. In both cases the model matches fairly well the data, even though the simulation slightly overestimates the educational outcomes of the lower quantiles. Even conditioning on unobserved heterogeneity observed cognitive skills are important schooling predictors.

Table 6 reports the wage equation estimated coefficients and standard errors. For computational reasons we divided the wages by a thousand, therefore all the coefficients in this table and in the following ones are also scaled by the same factor. The three constant terms show that type 1 ranks last with the lowest intercept, followed by type 3 and type 2 with the highest. The cognitive skills coefficients

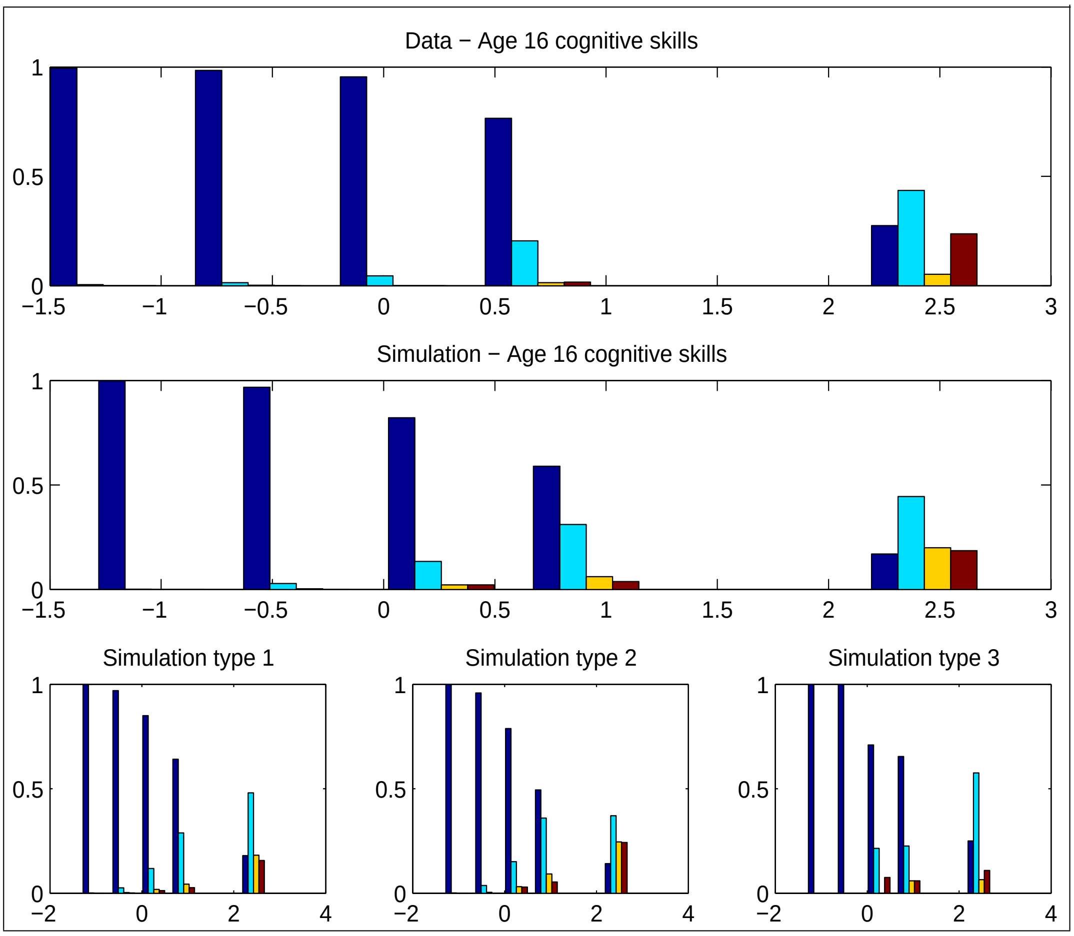

Figure 6: Qualifications by age 16 skills, true and simulated

are positive and highly significant as expected. The estimated return to schooling depends on the level of experience. With ten years of experience the return to O Levels is 5%, to A Levels is 13% and to Higher Education is 30%. These returns are lower than what we found by running a simple OLS. The difference is due to the selection on unobservables.

Table 7 reports the coefficients in the utility of schooling. The model suggests that individuals had a negative utility from O Levels, with the utility being the highest for type 2. The negative utility is due to the relatively high return to O Levels and the absence of foregone earnings for this choice. Given free education, the only reason not to take O Levels must be their relatively high effort cost. The utility from A Levels and Higher Education are instead positive because of foregone earnings and its not surprising because education was totally free.

The $\boldsymbol{\lambda}$ coefficients in table 8 indicate that the probability of success in obtaining a qualification once enrolled is increasing in both age 16 and age 7 cognitive skills. The difference across types do not appear to be statistically significant. Since we can not distinguish between enrolment and achievement in Higher Education we can not estimate λ_3. Nevertheless, rather than fixing $\lambda_3 = 1$, in the model we set $\lambda_3 = \lambda_2$. This should be a more realistic approximation of the true λ_3. [19]

The $\boldsymbol{\delta}$ coefficients in table 9 have the expected sign. We estimate an age 16 specific constant (δ_{01})

[19] Essentially we force individuals to believe there is a possibility of failing in Higher Education even though in the likelihood no one does.

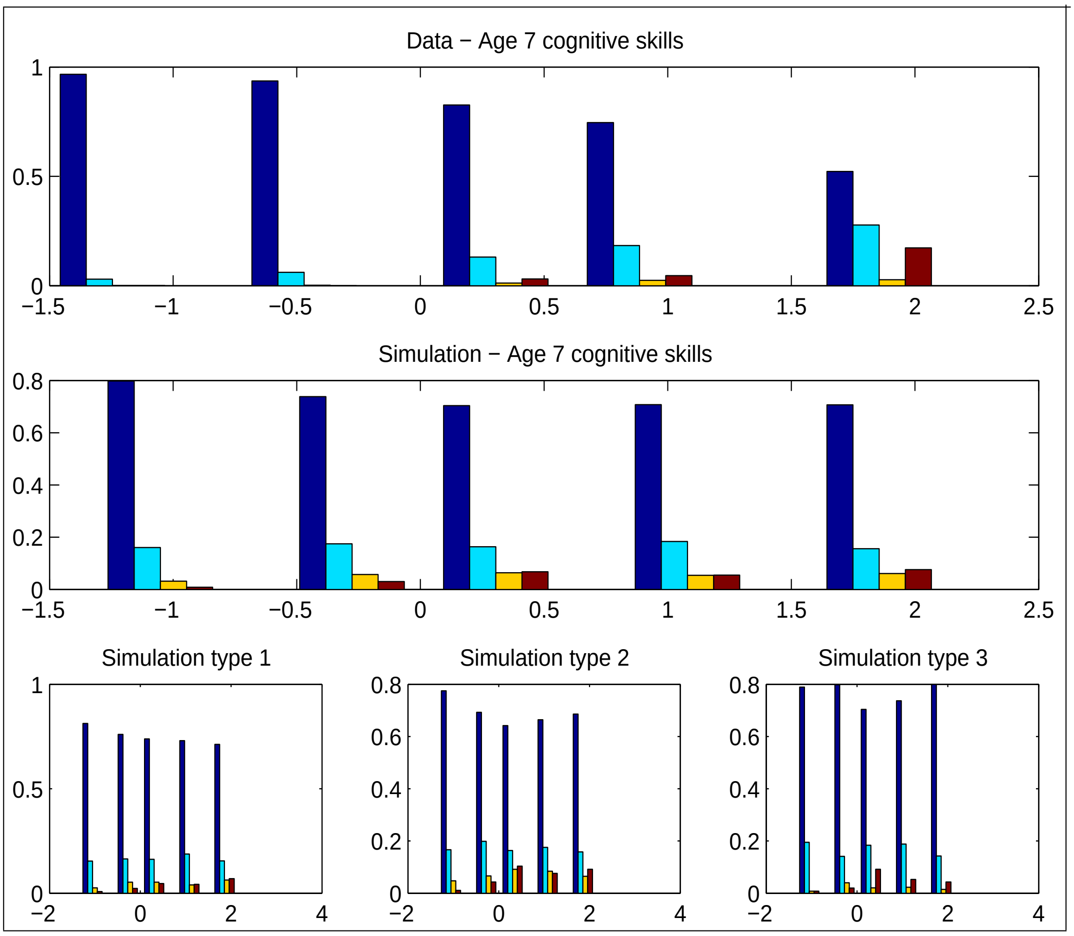

Figure 7: Qualifications by age 7 skills, true and simulated

and O Levels interaction (δ_{11}). From the data, we observed that those who entered the labor market immediately had a very large chance of finding a job, even with no academic qualification. Perhaps they already had a job or some link to a work place. The additional parameters correct for this. Type 3 individuals are the most likely to have a job offer, followed by type 2 and type 1.

The ϕ coefficients in table 10 show the expected sings. Being skilled, educated and experienced all reduce the probability of being fired. Again type 3 individuals are the least likely to be fired, followed by type 2 and type 1.

8.2 A Policy Experiment

As a policy experiment we simulate the effect of two alternative subsidies: a reduction in tuition fees, at any stage of the post 16 education process, or a subsidy to families when the individuals are still in their early childhood. The question we have in mind is which of the two would be more effective if we want to increase the education attainment of a new-born cohort or, given our data, a cohort that is still in the first decade of its life.

To start off, let us assume that the Government aims to increase Higher Education enrolment by 1% of the sample. We then use our model to compute the level of a tuition fees subsidy (grant) that would ensure such an increase. We alternate a Higher Education subsidy with an A Levels and O Levels ones.

With the exception of the last two columns, discussed later on, table 11 shows the result of performing this experiment. The upper part of the table reports the fraction of individuals enrolled in education at each age with or without subsidy. [20] We just report the results up to age 18 because this is when individuals enroll in higher education. The second part of the table reports respectively the amount of the subsidy, its ratio with the median parental income, and the per-capita cost. The latter takes into account the length and eligibility of each subsidy. An HE subsidy would be given for 3 years but only to those who enroll in HE. An AL subsidy would be given for 2 years to those who enroll in AL. An OL subsidy would be given once to those who enroll in OL.

Both an HE and AL subsidy succeeded in raising Higher Education enrolment by 1%, but the HE subsidy is more efficient. It is lower both in absolute terms and per-capita. The 1402 pounds subsidy correspond to 9.5% of the median parental income, costing 389 pounds each. We could not find an OL subsidy that would satisfy our initial requirement. The algorithm failed to find any improvement when it reached a subsidy worth more than 16824 pounds. For that value 86% of the individuals are enrolling in O Levels and yet the vast majority leave before Higher Education.

This result is clearly not surprising and embedded in the dynamic model. Subsidizing O Levels and A Levels does not change the incentives to get into Higher Education. The reason why some would enroll in Higher Education with OL and AL subsidies is that once the subsidy has brought them into O Levels or A Levels, they might get a large and positive taste shock for schooling and therefore stay even longer. Thus the main reason to test the effect of the grant at different educational levels is to show its overall effect. Even though these different subsidies are all raising HE enrolment by 1% (but in the OL case), they have different effects on AL and OL enrolment. As the second and third rows of table 11 show, the AL grant leads to higher AL enrolment than a HE grant. A similar argument applies to the OL grant. In defining the most efficient subsidy we are therefore assuming that the Government is only aiming for an increase in Higher Education while the externalities generated by higher OL and AL enrolment are negligible.

It is worth noting that Keane and Wolpin (1997) find a 8.4% increase in college graduation rates following a $ 2000 grant in a similar set up, a much bigger effect that we find. It is also true that pre-grant college graduation rate in their sample is 24.2%, also much larger than our 8.2%. It is hard to say whether this difference is due to differences in the educational systems and education incentives between the US and the UK, to different years of birth or to the different way we model the choices.

Next, we try to test whether a parental income subsidy would be more efficient than the HE one in pursuing our 1% increase. Our hypothesis is that a parental income subsidy would increase the age 16 cognitive skill endowment, and in turn enrolment in education. In figure 6 we have seen that our model predicts that highly skilled individuals are more likely to enroll. There is no other direct effect of a parental income subsidy on educational choices, not even through the unobserved heterogeneity. Although this might appear restrictive, it is in line with the current literature findings that educational choices are affected by parental income only indirectly through the stock of skills.

We assume that our age 16 skill level is a function of innate ability μ and a history of parental background inputs PB:

$$K_{16} = f(\mu, PB_1, .., PB_{16}) \tag{17}$$

in particular, we are interested in disentangling the effect of parental income among the parental background factors. If we assume f to be linear in its arguments then:

[20] Note that even without subsidy the figures in table 11 are different from those in table 4. This is because table 4 takes into account sample attrition: i.e. sample attrition is changing the distribution of skills and parental income over age in our sample, and the simulation in table 4 corrects for that. In table 11 instead we keep the initial distribution of skills and parental income at age 16^a to simulate the whole process.

$$E(K_{16}|Y_1, .., Y_{16}) = \boldsymbol{\eta}\boldsymbol{Y} \tag{18}$$

where $\boldsymbol{\eta}$ is a 1×16 vector of coefficients, and $\boldsymbol{Y}$ is a 16×1 vector of parental income inputs. Ideally we would like to know $max\{\boldsymbol{\eta}\}$ to subsidize parents at the most efficient point in time. We could think of a one off monetary subsidy, or voucher, that has to be spent within the year.

Unfortunately, our data do not provide us which such a long history of parental income: income is reported only at age 16. However, we have rich information on parental background characteristics at age 16, 11 and 7. These time varying variables would include father's social class, whether the mother was working, region of residence and of course age. Call $\boldsymbol{PBC}_t$ the vector including these variables at any time t.

We then try to infer parental income at a few points in time by first estimating a reduced form

$$Y_{16} = \boldsymbol{\zeta}_{16}\boldsymbol{PBC}_{16} + \upsilon \tag{19}$$

and, under the assumption that $\boldsymbol{\zeta}_{16} \approx \boldsymbol{\zeta}_{11} \approx \boldsymbol{\zeta}_7$, by computing

$$\hat{Y}_t = \hat{\boldsymbol{\zeta}}_{16}\boldsymbol{PBC}_t \tag{20}$$

for $t = 7, 11, 16$. Here $\hat{\boldsymbol{\zeta}}_{16}$ includes the constant and therefore the permanent income component. Finally we take

$$E(K_{16}|\hat{Y}_7, \hat{Y}_{11}, \hat{Y}_{16}, K_7) = \boldsymbol{\eta}\hat{\boldsymbol{Y}} + \varsigma K_7 \tag{21}$$

where we include K_7 as a proxy of innate ability. A simple OLS estimation gives:

$$\eta_7 = 0.01; \quad \eta_{11} = 0.026; \quad \eta_{16} = 0.066; \tag{22}$$

all being statistically significant. [21]

There is no need to say that these estimates are quite rough approximations of the true $\boldsymbol{\eta}$'s. Upward bias is likely to be induced by omitted variables and downward bias by measurement error. The economic literature on the topic is still far from being able to clearly establish the true link between parental income and skills. [22] Dahl and Lochner (2005) use the Earning Tax Credit scheme as an instrument for parental income. They estimate that a 10000 dollars increase in income raises math test scores by 21% and reading test scores by 36% of a standard deviation. Their estimated is above what usually found in the literature using other instruments, fixed effects or simple OLS. [23] Our results suggest that an increase of 10000 pounds in 2001 prices (around 15000 dollars), would push up the principal component by 50% of a standard deviation, close to their result. In any case, we do not feel in the position of claiming a consistent estimate of $\boldsymbol{\eta}$ either. We only have limited information on parental income, and we do not believe any of the variables at our disposal could be a valid instrument.

Therefore we use the model to compute the value of η for which a one-off parental income subsidy would be as powerful as the HE tuition fees subsidy. Given that the two subsidies would have to cost the same, we set the parental income subsidy to 389 pounds, the per-capita cost of the HE subsidy. The result is reported in the fourth column of table 11 (P. Inc. Sub (1)). We do manage to get an increase in HE enrolment by 1%, but this occurs for a value of $\tilde{\eta}$ equal to 0.216. This would be 3.24 times our estimated η_{16} implying that an increase of 10000 pounds would push up the score by 168% of a standard

[21] Income is in 1000 pounds.

[22] See Todd and Wolpin (2003), Cunha and Heckman (2006) and Cunha, Heckman, Lochner, and Masterov (2005) for a discussion on the skills production function.

[23] Their data are the children of the NLSY79 cohort, aged 9.5 on average.

deviation. This is far above any estimate found in the skill literature, making it unlikely that even with a consistent estimate of η the parental income subsidy could have been as powerful as the HE one.

It is instructive to compare the age 16 cognitive skills distribution before and after the parental income subsidy. The mean skill after the parental subsidy is 6.5% of a standard deviation larger then before the subsidy. This implies that any policy aiming to achieve our +1% in enrolment only through cognitive skill accumulation would have to increase skills by at least 6.5% of a standard deviation.

Given the non-linearity of the schooling choice model, this result depends inevitably on the targeted raise in Higher Education enrolment. The ratio between $\tilde{\eta}$ and our estimated η_{16} monotonically decreases as we target a larger augment in Higher Education enrolment, going down to 2.15 when we move from a +1% to a +5% enrolment. However such an increase would likely have a strong effect on the return to education and cognitive skills, which our model does not account for. Heckman, Lochner, and Taber (1998) and Lee (2005) predicts a smaller impact of tuition subsidies once the general equilibrium effect, i.e. a lower return to schooling, is considered. Intuitively, if a general equilibrium effect was to appear in our model, the result of our comparison should be reinforced. While both kinds of subsidy would lead to a lower return to schooling, shifting the distribution of cognitive skills towards the right would also lower its return and therefore weaken the positive link between cognitive skill and education. Thus, a larger increase in cognitive skills would be needed to reach the targeted increase in Higher Education enrolment.

Next, we check what would happen if rather than targeting the age 16 cognitive skills stock, we were to target the age 7 one. If cognitive skills are shaped early in life, and if age 7 cognitive skills have an effect on schooling choices even conditioning on age 16 ones, as the coefficients indicate, than targeting the age 7 skills stock might be more efficient. This assumes that K_7 is not exogenous and therefore not a good proxy of innate ability. The ς coefficient in the production function, equation (21), would be inconsistent and likely upward biased. Nevertheless, keeping this in mind, it is interesting to ask what change in age 7 cognitive skills is needed to match the effect of a 389 pounds Higher Education subsidy. We actually fall below the target by 0.86 percentage points since the effect of age 7 skills on educational choices is not strong enough. However, when the algorithm stops the return to parental income is already 0.21, that is 3.15 our OLS estimate. If $\hat{\varsigma}$ is upward biased then $\tilde{\eta}$ is a lower bound.

We also try to set a +1% target for A Level and O Level achievement. The Government might just value raising education to post-compulsory levels without necessarily increasing Higher Education enrolment. The results are shown in tables 12 and 13. As we would expect, the best way to reach the target is to give a grant at the target level, that is subsidize A Levels if the target is +1% enrolment in A Level, or similarly for O Levels. The parental income subsidy still require very high returns to be efficient, respectively 3.03 (A Levels) and 2.40 (O levels) times our OLS estimate when targeting age 16 skills.

A few more points worth noting. One could argue that this policy experiment is not very informative given that this cohort was born in 1958 and participation in Higher Education is much higher for the current cohorts. The importance of cognitive skills might have changed over time. Yet higher participation, either due to larger supply or demand, is likely to decrease the importance of cognitive skills, given that a now less elitarian group of individuals staying on. Therefore, our model is likely to over-estimate the impact of cognitive skills and, as a consequence, of the parental income subsidy, strengthening our conclusions.

The linear utility implies no borrowing constraints and no risk aversion. If the individuals did face a borrowing constraint then tuition fees subsidies (grants) would be even more effective than our model predicts. Concerning risk aversion, the only source of randomness in our model is the taste shock in the utility of schooling. There is no a priori reason why a tuition subsidy or an increase in cognitive skills should have very different effects due to risk aversion.

9 Conclusions

In this paper we build a dynamic structural model of educational choices and exploit rich cohort data to investigate the importance of observed cognitive skills, financial incentives and unobserved heterogeneity in educational choices. The model predicts that highly skilled individuals stay longer in education and it does well in replicating the choices conditional on the skill group.

We then simulate the effect of two educational subsidies equal in cost but different in timing. The first consists of grants assigned directly to the individuals aged between 16 and 18. The second is a subsidy assigned to the parents earlier on, when the cohort is still in its childhood. Our aim is to test whether an unconditional parental subsidy, with its indirect effect on cognitive skills, could be more efficient in fostering educational enrolment than a direct educational grant. When we target an increase in Higher Education enrolment equal to 1% of the population, our results suggest that this is not the case. Unless the effect of parental income on skills is implausibly high, a direct grant in the form of a tuition subsidy seems more efficient even in the absence of short term financial constraints. Although cognitive skills accumulated during childhood play a key role in the educational decisions, an unconditional financial subsidy to parents is not the best policy. This result is robust even if we target a different increase in Higher Education enrolment or if we target an increase at a lower education level.

The overall result does not call a halt to additional investment in cognitive skill accumulation during childhood. We only consider the link between cognitive skills and parental income, surely weaker than between skills and parental investment, such as child care or parental time. In our opinion, the insight of the paper is the opposite, cognitive skills are important but changes in educational choices demand a sizeable shift in the cognitive skills distribution.

References

J. Arcidiacono. "Affirmative Action in Higher Education: How do Admission and Financial Aid Rules Affect Future Earnings?". *Econometrica*, 73(5):1477–1524, 2005.

J. Arcidiacono and J. B. Jones. "Finite Mixture Distributions, Sequential Likelihood, and the EM Algorithm". *Econometrica*, 71(2):933–946, 2003.

O. Attanasio, E. Fitzsimons, and C. Meghir. "The Impact of a Conditional Education Subsidy on the Activities of Young Adults in the UK", 2004. mimeo.

C. Belzil and J Hansen. "Unobserved Ability and the Return to Schooling". *Econometrica*, 70(2): 2078–2091, 2002.

S. Cameron and J. Heckman. "The Dynamics of Educational Attainment for Black, Hispanic and White Males". *The Journal of Political Economy*, 109(3):455–499, 2001.

S. Cameron and J. Heckman. "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males". *The Journal of Political Economy*, 106(2):262–333, 1998.

S. V. Cameron and C. Taber. "Estimation of Educational Borrowing Constraints Using Returns to Schooling". *Journal of Political Economy*, 112(1):132–182, 2004.

P. Carneiro and J. Heckman. "The Evidence on Credit Constraints in Post-Secondary Schooling". *The Economic Journal*, 112:705–734, 2002.

G. Conlon. "The differential in earnings premia between academically and vocationally trained males in the United Kingdom", 2001. Centre for the Economics of Education Discussion Papers.

F Cunha and J. Heckman. "Formulating, Identifying, and Estimating the Technology for the Formation of Skills". *Journal of Human Resources*, (forthcoming), 2006.

F. Cunha, J. Heckman, L. Lochner, and D. Masterov. "Interpreting the Evidence on Life Cycle Skill Formation". *NBER working papers*, (11331), 2005.

G. B. Dahl and L. Lochner. "The Impact of Family Income on Child Achievement". *NBER working papers*, (11279), 2005.

L. Dearden, L. McGranahan, and B. Sianesi. "The Role of Credit Constraints in Educational Choices: Evidence from the NCDS and BCS70". Centre For the Economics of Education discussion paper, 2004. London School of Economics, London.

Z. Eckstein and K. Wolpin. "The Specification and Estimation of Dynamic Stochastic Discrete Choice Models: A Survey". *The Journal of Human Resources*, 24(4):562–598, 1989.

B. S. Everitt and D. J. Hand. *Finite Mixture Distributions*. Chapman and Hall, London and New York, 1981.

J. Heckman and B. Singer. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data". *Econometrica*, 52(2):271–320, 1984.

J. J. Heckman, L. Lochner, and C. Taber. "General-Equilibrium Treatment Effects: A Study of Tuition Policy". *American Economic Review*, 88(2):381–386, 1998.

J. J. Heckman, J Stixrud, and S. Urzua. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior". *Journal of Labor Economics*, 24(3):411–482, 2006.

R. J. Herrnstein and C. Murray. *The Bell Curve*. The Free Press, 1994.

H. Ichimura and C. Taber. "Semiparametric Reduced-Form Estimation of Tuition Subsidies". *American Economic Review*, 92(2):286–292, 2002.

M. Keane and K. Wolpin. "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment". *International Economic Review*, 42(4):1051–1103, 2001.

M. Keane and K. Wolpin. "The Career Decisions of Young Men". *The Journal of Political Economy*, 105(3):473–522, 1997.

Donghoon Lee. "An Estimable Dynamic General Equilibrium Model of Work, Schooling and Occupational Choice". *International Economic Review*, 46(1):1–34, 2005.

T. Magnac and D. Thesmar. "Identifying Dynamic Discrete Decision Processes". *Econometrica*, 70(2): 801–816, 2002.

G. McLachlan and D. Peel. *Finite Mixture Models*. John Wiley and Sons, INC., 2000.

A. Roy. "Some Thoughts on the Distribution of Earnings". *Oxford Economics Papers*, 3:135–146, 1951.

J. Rust. *"Structural Estimation of Markow Decision Processes"*, volume IV of *Handbook of Econometrics*. Elsevier Science, 1994. editors R.F. Engle and D.L. McFaddden.

P. Todd and K. Wolpin. "On the Specification and Estimation of the Production Function for Cognitive Achievement". *The Economic Journal*, 113:3–33, 2003.

R. J. Willis and S. Rosen. "Education and Self-Selection". *The Journal of Political Economy*, 87:S7–S36, 1979.

Table 1: Status by age (Number of observations)

Age	Education	Work	Unemployment	Total
16[a]	1,811	0	1,063	2,874
	(63.01)	(0.00)	(36.99)	(100.00)
16[b]	535	2,178	161	2,874
	(18.62)	(75.78)	(5.60)	(100.00)
17	535	2,222	109	2,866
	(18.67)	(77.53)	(3.80)	(100.00)
18	197	2,349	153	2,699
	(7.30)	(87.03)	(5.67)	(100.00)
19	197	2,341	129	2,667
	(7.39)	(87.78)	(4.84)	(100.00)
20	197	2,318	129	2,644
	(7.45)	(87.67)	(4.88)	(100.00)
21	0	2,407	153	2,560
	(0.00)	(94.02)	(5.98)	(100.00)
22	0	2,208	267	2,475
	(0.00)	(89.21)	(10.79)	(100.00)
23	0	2,190	252	2,442
	(0.00)	(89.68)	(10.32)	(100.00)
24	0	1,729	120	1,849
	(0.00)	(93.51)	(6.49)	(100.00)
25	0	1,714	119	1,833
	(0.00)	(93.51)	(6.49)	(100.00)
26	0	1,692	123	1,815
	(0.00)	(93.22)	(6.78)	(100.00)
27	0	1,687	113	1,800
	(0.00)	(93.72)	(6.28)	(100.00)
28	0	1,692	100	1,792
	(0.00)	(94.42)	(5.58)	(100.00)
29	0	1,686	92	1,778
	(0.00)	(94.83)	(5.17)	(100.00)
30	0	1,686	81	1,767
	(0.00)	(95.42)	(4.58)	(100.00)
31	0	1,676	83	1,759
	(0.00)	(95.28)	(4.72)	(100.00)
32	0	1,654	101	1,755
	(0.00)	(94.25)	(5.75)	(100.00)
33	0	1,458	81	1,539
	(0.00)	(94.74)	(5.26)	(100.00)
34	0	1,413	78	1,491
	(0.00)	(94.77)	(5.23)	(100.00)
35	0	1,407	76	1,483
	(0.00)	(94.88)	(5.12)	(100.00)
36	0	1,400	78	1,478
	(0.00)	(94.72)	(5.28)	(100.00)
37	0	1,391	82	1,473
	(0.00)	(94.43)	(5.57)	(100.00)
38	0	1,381	85	1,466
	(0.00)	(94.20)	(5.80)	(100.00)
39	0	1,372	87	1,459
	(0.00)	(94.04)	(5.96)	(100.00)
40	0	1,360	90	1,450
	(0.00)	(93.79)	(6.21)	(100.00)
41	0	1,332	108	1,440
	(0.00)	(92.50)	(7.50)	(100.00)
Total	4,047	45,943	4,251	54,241
	(7.46)	(84.70)	(7.84)	(100.00)

Table 2: Wage (Sample averages)

	age=23	age=33	age=41
Wage	12326.47	19175.00	23904.71
	(3220.47)	(7209.19)	(11409.47)
	Wage by qualification.		
Wage — No Qualif.	12094.48	17607.02	21381.15
	(3233.81)	(6115.98)	(9011.91)
Wage — O Levels	13125.45	22878.02	28981.42
	(3050.66)	(7750.02)	(12156.04)
Wage — A Levels	13415.03	25742.56	37967.17
	(2670.46)	(6602.89)	(18981.43)
Wage — H.E.	13403.48	28595.93	38780.83
	(3022.19)	(7964.21)	(15630.93)

Standard Deviations in brackets.
Yearly wages are derived by first computing hourly wages and then multiplying them by 52. Wages are in January 2001 prices.

Table 3: Cognitive Skills and Parental Income (Sample averages)

	age 16	age 7
Math Score	13.68	5.33
	(7.26)	(2.44)
Reading Score	25.70	23.15
	(7.10)	(7.00)
Principal Component	-2.28e-10	-2.34e-09
	(1.28)	(1.17)
Parental Income	16226.79	—-
	(6128.49)	(—-)

Standard Deviations in brackets.

Table 4: Model Fit - Status (Proportions)

	Education		Work		Unemp.	
Age	True	Predicted	True	Predicted	True	Predicted
16[a]	0.6301	0.6241	0	0	0.3699	0.3759
16[b]	0.1862	0.1688	0.7578	0.7805	0.0560	0.0507
17	0.1867	0.1675	0.7753	0.7850	0.0380	0.0475
18	0.0730	0.0701	0.8703	0.8688	0.0567	0.0611
19	0.0739	0.0672	0.8778	0.8794	0.0484	0.0534
20	0.0745	0.0657	0.8767	0.8798	0.0488	0.0545
21	0	0	0.9402	0.9442	0.0598	0.0558
22	0	0	0.8921	0.9436	0.1079	0.0564
23	0	0	0.8968	0.9418	0.1032	0.0582
24	0	0	0.9351	0.9422	0.0649	0.0578
25	0	0	0.9351	0.9431	0.0649	0.0569
26	0	0	0.9322	0.9426	0.0678	0.0574
27	0	0	0.9372	0.9430	0.0628	0.0570
28	0	0	0.9442	0.9437	0.0558	0.0563
29	0	0	0.9483	0.9454	0.0517	0.0546
30	0	0	0.9542	0.9451	0.0458	0.0549
31	0	0	0.9528	0.9452	0.0472	0.0548
32	0	0	0.9425	0.9448	0.0575	0.0552
33	0	0	0.9474	0.9465	0.0526	0.0535
34	0	0	0.9477	0.9473	0.0523	0.0527
35	0	0	0.9488	0.9475	0.0512	0.0525
36	0	0	0.9472	0.9481	0.0528	0.0519
37	0	0	0.9443	0.9490	0.0557	0.0510
38	0	0	0.9420	0.9484	0.0580	0.0516
39	0	0	0.9404	0.9471	0.0596	0.0529
40	0	0	0.9379	0.9460	0.0621	0.0540
41	0	0	0.9250	0.9443	0.0750	0.0557

Table 5: Model Fit - Wage (Sample average)

	True	Predicted	True	Predicted	True	Predicted	True	Predicted
	Overall		Age 23		Age 33		Age 41	
	17750	18003	12326	12635	19175	18094	23905	23518
No Qualif	16381	16590	12094	12338	17607	16749	21381	20866
O Levels	20987	20722	13125	13463	22878	20756	28981	28299
A Levels	25297	22413	13415	13061	25743	22089	37967	32532
HE	26341	25299	13403	13797	28596	24786	38781	37809

Table 6: Estimated Coefficients - Wage Equation

	Estimate	S.E.	ratio
σ_u	0.2789	(0.0027)	(104.5020)
Type 1	2.2286	(0.0147)	(151.1216)
Type 2 deviation from type 1	0.3229	(0.0118)	(27.3637)
Type 3 deviation from type 1	0.0249	(0.0644)	(0.3862)
α_{k2}	0.0654	(0.0073)	(8.9546)
α_{k1}	0.0190	(0.0056)	(3.4003)
α_{q1}	-0.0614	(0.0356)	(-1.7277)
α_{q2}	-0.0676	(0.1227)	(-0.5511)
α_{q3}	0.0354	(0.0559)	(0.6331)
α_x	0.0278	(6.7610e-004)	(41.0976)
α_{x1}	0.0118	(0.0017)	(6.7742)
α_{x2}	0.0202	(0.0060)	(3.3557)
α_{x3}	0.0271	(0.0030)	(9.0672)

Table 7: Estimated Coefficients - Utility of Schooling

	Estimate	S.E.	ratio
σ_{ϵ_1}	18.0527	(2.5486)	(7.0834)
σ_{ϵ_2}	14.9309	(4.1329)	(3.6127)
σ_{ϵ_3}	21.9125	(9.3737)	(2.3376)
Type 1	-5.5077	(1.1841)	(-4.6512)
Type 2 deviation from type 1	4.1302	(1.6581)	(2.4909)
Type 3 deviation from type 1	2.6691	(2.4009)	(1.1117)
γ_{AL}	12.0880	(3.2117)	(3.7637)
γ_{HE}	11.8215	(3.3493)	(3.5295)

Table 8: Estimated Coefficients - Probability of Success in School

	Estimate	S.E.	ratio
Type 1	-0.9936	(0.0816)	(-12.1709)
Type 2 deviation from type 1	0.0754	(0.1477)	(0.5108)
Type 3 deviation from type 1	-0.0412	(0.2444)	(-0.1687)
λ_{1k2}	1.3638	(0.0371)	(36.7728)
λ_{1k1}	0.1081	(0.0382)	(2.8314)
λ_{AL}	0.5481	(0.1631)	(3.3613)
λ_{2k2}	0.4946	(0.0857)	(5.7715)
λ_{2k1}	0.1357	(0.0387)	(3.5074)

Table 9: Estimated Coefficients - Probability of having a Job Offer

	Estimate	S.E.	ratio
Type 1	0.4841	(0.0674)	(7.1855)
Type 2 deviation from type 1	0.2103	(0.1264)	(1.6633)
Type 3 deviation from type 1	1.3726	(0.0599)	(22.9256)
δ_{k2}	0.1211	(0.0305)	(3.9739)
δ_{k1}	0.0054	(0.0258)	(0.2090)
δ_{q1}	0.0986	(0.0911)	(1.0829)
δ_{q2}	0.0229	(0.2003)	(0.1145)
δ_{q3}	0.9012	(0.1520)	(5.9292)
δ_x	-0.0741	(0.0050)	(-14.8447)
δ_{01}	1.2630	(0.0726)	(17.4067)
δ_{11}	-0.7956	(0.1354)	(-5.8780)

Table 10: Estimated Coefficients - Probability of being fired

	Estimate	S.E.	ratio
Type 1	-1.7899	(0.0335)	(-53.4109)
Type 2 deviation from type 1	-0.6224	(0.0836)	(-7.4437)
Type 3 deviation from type 1	-1.0765	(0.0778)	(-13.8312)
ϕ_{k2}	-0.1378	(0.0198)	(-6.9570)
ϕ_{k1}	0.0034	(0.0159)	(0.2120)
ϕ_{q1}	-0.1980	(0.0588)	(-3.3684)
ϕ_{q2}	0.0741	(0.1616)	(0.4585)
ϕ_{q3}	0.0526	(0.1138)	(0.4619)
ϕ_x	-0.0200	(0.0024)	(-8.4790)

Table 11: Policy Experiment - +1% in HE enrolment

	No Sub	O Levels Sub.	A Levels Sub.	HE Sub	P. Inc. Sub (1)	P. Inc. Sub (2)
16^a	0.6241	0.8611	0.6278	0.6255	0.6363	0.6326
16^b	0.1642	0.1747	0.1876	0.1717	0.1773	0.1712
17	0.1642	0.1747	0.1876	0.1717	0.1773	0.1712
18	0.0825	0.0852	0.0925	0.0925	0.0925	0.0839
Subsidy	0	16824	1932	1402	1402	1402
Subsidy/Median Income	0	1.1306	0.129	0.094	0.094	0.094
Per-Capita Cost	0	14487	725	389	389	389

Table 12: Policy Experiment - +1% in AL enrolment

	No Sub	O Levels Sub.	A Levels Sub.	HE Sub	P. Inc. Sub (1)	P. Inc. Sub (2)
16^a	0.6241	0.8779	0.6259	0.6256	0.6328	0.6305
16^b	0.1642	0.1742	0.1742	0.1742	0.1742	0.1724
17	0.1642	0.1742	0.1742	0.1742	0.1742	0.1724
18	0.0825	0.0849	0.0870	0.0949	0.0912	0.0866
Subsidy	0	18529	877	1724	877	877
Subsidy/Median Income	0	1.245	0.059	0.115	0.059	0.059
Per-Capita Cost	0	16267	305	490	305	305

Table 13: Policy Experiment - +1% in OL enrolment

	No Sub	O Levels Sub.	A Levels Sub.	HE Sub	P. Inc. Sub (1)	P. Inc. Sub (2)
16^a	0.6241	0.6341	0.6341	0.6341	0.6341	0.6335
16^b	0.1642	0.1653	0.2158	0.2187	0.1747	0.1672
17	0.1642	0.1653	0.2158	0.2187	0.1747	0.1672
18	0.0825	0.0827	0.1054	0.1423	0.0863	0.0857
Subsidy	0	689	4319	8563	689	689
Subsidy/Median Income	0	0.046	0.290	0.575	0.046	0.046
Per-Capita Cost	0	437	1864	3655	437	437