Returns to Education for the 'Marginal Learner': Evidence from the BCS70

Lorraine Dearden

Leslie McGranahan

Barbara Sianesi

December 2004

Published by
Centre for the Economics of Education
London School of Economics
Houghton Street
London WC2A 2AE

ISBN 07530 1739 3
Individual copy price: £5

The Centre for the Economics of Education is an independent research centre funded by the Department for Education and Skills. The views expressed in this work are those of the author and do not reflect the views of the DfES. All errors and omissions remain the authors.

Executive Summary

To inform educational policy, it is often more relevant to know the returns to educational investments that accrue (or would have accrued) to the marginal entrant in education, rather than to the average learner.

In this paper, we focus on two key educational choices and estimate the corresponding individual wage returns for variously defined marginal learners, in particular:

1. the returns to staying on in full-time education past the age of 16 for the marginal stayer;
2. the returns to completing any form of higher education (HE) compared to at least level 2 for the marginal entrant in HE who holds at least level 2 qualifications.
   Where appropriate, we contrast these results to those for the returns to HE compared to anything less.

The key issue naturally concerns the identification of the 'marginal learner' for the educational investments we consider. By definition, it is the individual at the margin in that educational decision, but how can one identify those who are at the margin?

Since we do not have a well-defined policy instrument which would allow us to define the so-called Marginal Treatment Effect developed in the recent evaluation literature, in this work we look at the average return for individuals at the margin in the educational decision, where 'margin' is defined in various alternative ways. We feel that all of these 'marginal' individuals are of interest in their own right, in fact allowing us to shed light on distinct questions. In particular, we explore the following characterisations of the marginal learner for a given educational level:

1. those who have achieved the level of education being looked at – the corresponding return is the Average Effect of Treatment on the Treated (ATT);
2. those who could have but did not achieve that level of education – the corresponding return is the Average Effect of Treatment on the Non-Treated (ATNT);
3. those eligible to undertake the qualification, irrespective of whether they actually achieved it or not – the corresponding return is the Average Treatment Effect (ATE);
4. those with low (medium or high) values of the probability of achieving that educational outcome – the corresponding return is the Average Treatment Effect for individuals whose probability of achieving education falls within a given interval;
5. those defined as marginal entrants on a policy basis, in particular groups defined in terms of their ability, socio-economic background or family income – the corresponding return, depending on how it is calculated, is the ATT, the ATNT or the ATE for the target individuals.

To control for selection into education, we rely on the extensive individual and family background information available in the 1970 British Cohort Study and estimate wage returns to staying on and to HE enjoyed by variously defined 'marginal' groups when aged 29/30.

Our general finding is that there are substantial returns to both staying on and to HE for all subgroups of the population, though we have uncovered some variation in returns, especially for men. Returns are invariably found to be higher for women than for men – though as we

argue, this ranking needs to be interpreted with caution.

As to staying on past compulsory schooling:

– Our estimates of the returns to staying on are around 11% for men and 18% for women.

– For either males of females, we have not uncovered statistically significant heterogeneity in returns by actual staying-on decision, staying-on probability, social class, family income or ability.

– The lowest returns overall (6-8%) are those that male drop-outs of either low-ability or low social class would have enjoyed had they stayed on.

– For both mean and women, individuals from a low-income family who drop out would have enjoyed substantial returns from staying on (around 13% for men and 17% for women). This result should be interpreted with care given the patterns of missing income information. If we take the additional caution required when extrapolating results for the mid 80s to a more recent period, this finding is indicative that individuals from poorer backgrounds who do not stay in school could considerably benefit were they encouraged to do so.

As to the attainment of any form of higher education conditional on having achieved at least level 2 qualifications:

– We find sizeable average wage returns to HE relative to holding at least a level 2 qualification, estimates being around 15% for men and 22% for women. (Returns to HE compared to anything less are obviously larger, around 20% for men and 25% for women).
In fact, nearly all of our estimates of returns to HE are higher than the returns to staying on for the same subgroups.

– In terms of returns by attainment probability, we find that it is those men and women who are 'indifferent' (i.e. in the 25-50% probability range) between undertaking HE or remaining at level 2 or 3 who experience the highest returns from moving on to HE.

– For men, we find that the returns to HE are substantially higher for the more disadvantaged groups. In particular, low socio-economic class men enjoy significantly higher returns than high social class men; men from a low-income family have significantly higher returns than those from a more well-off family; and low-ability men have slightly higher returns to HE than high-ability men.

– Among low-income men, we find higher returns to HE for those who did achieve HE than for those who stopped at level 2 or 3, which might be taken as an indication of informed, efficient sorting into HE.

– By contrast, for high-income males, as well as for men from high social class, the average HE return the non-treated would have enjoyed exceeds the average return enjoyed by the treated, this difference being statistically significant. One possible explanation relates to an expanded role of non-economic factors in the decision-making of the well-off groups.

– For the low social-class male sample, the average returns to HE for the treated and for the

non-treated are roughly equivalent.

– In sharp contrast to men, for women we find very similar returns to HE (compared both to at least level 2 and to anything less) across ability, income and social class groups, as well as for each treated and non-treated subgroups. This may be due to factors concerning selection into employment.

Returns to Education for the 'Marginal Learner': Evidence from the BCS70

Lorraine Dearden
Leslie McGranahan
Barbara Sianesi

1. Introduction 1

2. Who is the 'Marginal Learner'? 2
   - I. The treated 3
   - II. The non-treated 3
   - III. The eligibles 3
   - IV. Individuals with different attainment probabilities 3
   - V. Defined a priori on a policy basis 4

3. Methodology 5

4. Results 7
   - 4.1 Staying on past compulsory schooling 8
   - 4.2 Higher education 13

5. Conclusions 19

6. References 20

Appendices 21
- A – Methodological approach to estimate returns to the marginal learner 21
- B – Returns by propensity score bands 25
- C – Wage returns to HE compared to anything less 29
- D – Individuals with missing parental income at 16 31

Acknowledgements

The authors would like to thank Bob Butcher, Mutsa Chironga, John Elliot, Mark Franks, Karen Hancock, Paul Johnson and participants at the IFS seminar for comments on earlier versions of this work.

Lorraine Dearden is Director of the Centre for Early Years and Education Research at the Institute for Fiscal Studies. Leslie McGranahan and Barbara Sianesi are Senior Research Economists in the Education, Employment and Evaluation Sector at the Institute for Fiscal Studies.

1. Introduction

To help guide and direct policy, it is marginal rather than average returns to education that matter most. In deciding where to invest the extra pound, efficiency considerations require a comparison of the marginal benefit (or return) from the different educational investments. Note that the marginal benefit should include both the private and social net gains. Also, to decide about if and how to intervene, the policymaker should consider whether there are any market failures causing the individual to under-invest in education from a social point of view. Finally, equity concerns should be addressed separately, and could of course override any such calculation.

In this paper, we focus on an essential ingredient for answering this important question: what are the individual wage returns to education for the marginal learner?[1]

We focus on the returns to two key educational choices, and estimate:

- the returns to staying on post-compulsory schooling for the marginal stayer;
- the returns to completing any form of higher education (HE) compared to at least level 2 for the marginal entrant in HE who holds at least level 2 qualifications. Where appropriate, we contrast these results to those for the returns to HE compared to anything less.

The key issue naturally concerns the identification of the 'marginal learner' for the educational investments we consider, and this is discussed in the next section.

The most recent data we can use to adequately address these questions is the 1970 British Cohort Study (BCS70). This study, which has followed all individuals born between 5 April and 11 April 1970 since birth, contains the detailed information required for our analysis, in particular on ability, family background, education and employment. We evaluate the wage returns to our two educational choices for variously defined marginal learners when they were aged 29-30 in 1999/2000.

Despite our exclusive focus on the most recent available cohort, it has to be borne in mind that these individuals were taking their staying on decisions in 1986 and their HE participation decisions in 1989. The 'marginal' student, however defined, is very likely to have changed in profile since then, particularly in light of the massive expansion in HE participation which has taken place since the 1980s.

The even earlier birth cohort (the 1958 NCDS) would be even less representative of the 'marginal' student of today, and simply comparing returns between the two cohorts could be misleading.[2] For these reasons we do not report estimates of the returns to the marginal learner for the NCDS.

[1] In a companion paper, we have considered the extent to which individuals' educational choices as to staying on and to HE attainment are restricted by credit constraints (Dearden, McGrahanan and Sianesi, 2004).

[2] A comparison of the returns for 33-year-olds in 1991 (NCDS) to those for 30-year-olds in 1999/200 (BCS70) will conflate time effects, age effects, and cohort effects. As to time effects, the NCDS survey occurred in a recession year – in 1991 real GDP fell by 1.4%. By contrast, in the year of the BCS70 survey (1999/2000), real GDP grew by 3.8%. Individuals who work, and thus are observed, during a recession may be selected differently than individuals who work during an expansion, while educational qualifications may be rewarded differently in a recession than in an economic expansion. As to age effects, returns may vary over the life-cycle. The 3-year age difference between the two cohorts might be especially problematic for women, given that the age at first birth has been increasing and fertility declining between the two cohorts. Seventy-five percent of the 33 year-old

2. Who is the 'Marginal Learner'?

Who exactly is the 'marginal learner' for a given educational investment, and what characteristics define him or her?
By definition, it is the individual at the margin in that educational decision, but how can one identify those who are at the margin?

One definition of this 'margin' – and in fact the one on which the definition of the so-called Marginal Treatment Effect (MTE) is based in the recent evaluation literature – is to define the margin in terms of a *policy instrument* that affects educational participation. Formally, if we were interested in, say, HE and had a policy instrument, for example a tuition subsidy of x pounds, the MTE of HE is defined as the average return for those individuals who are indifferent between participating and not participating in HE at the given value x of the policy instrument. The MTE is thus the parameter required for evaluating the effect of a marginal change in policy (from x to $x+\varepsilon$ or $x-\varepsilon$) on the persons induced into (or out of) education by the change.

Since however we do not have a well-defined policy instrument, in this work we look at the average return for individuals at the margin in the educational decision, where 'margin' is defined in various alternative ways. We feel that all of these 'marginal' individuals are of interest in their own right, in fact allowing us to shed light on distinct questions. In particular, we suggest the following characterisations of the marginal learner for a given educational level:

1. those who have achieved the level of education being looked at – the corresponding return is the Average Effect of Treatment on the Treated (ATT);
2. those who could have but did not achieve that level of education – the corresponding return is the Average Effect of Treatment on the Non-Treated (ATNT);
3. those eligible to undertake the qualification, irrespective of whether they actually achieved it or not – the corresponding return is the Average Treatment Effect (ATE);
4. those with low (medium or high) values of the probability of achieving that educational outcome – the corresponding return is the Average Treatment Effect for individuals whose probability of achieving education falls within a given interval;
5. those defined as marginal entrants on a policy basis, e.g. high-ability individuals from a low-income family, or individuals from a low social parental class – the corresponding return, depending on how it is calculated, is the ATT, the ATNT or the ATE for the target individuals.

We now consider these alternatives in turn.

---

women in the NCDS had given birth to a child, as compared to 54% of the 30 year-old women in the BCS70. This in part due to the fact that the NCDS women are older when they are surveyed, but also due to increased fertility among the NCDS women until age 30; in fact, 66% of the NCDS women had given birth by age 30. Women with children have different work patterns than women without children, being far less likely to work and when they do, working for fewer hours per week. The two sub-samples of working women are thus most likely to have a different composition, which makes the interpretation of between-cohort changes in the returns for women who are observed in work a rather complex exercise. Finally, cohort effects may be reflected in changes in returns to qualifications which arise from cohort differences in human capital acquisition both in a demand and supply framework and in terms of students' composition. The latter is likely to be particularly important when trying to compare returns to 'marginal' learners between the two cohorts, in that the expansion of education and the changes in social and individual attitudes and aspirations will most likely have resulted in substantial changes in the composition of learners in general and of learners at the margin in particular.

I. ***Marginal learners*: the treated**

The 'treated' are those individuals who actually did choose to acquire the qualification of interest. The average return for them is the Average Effect of Treatment on the Treated (ATT). This is the parameter of interest if achievement of the qualification is voluntary and we wish to estimate the average return among those individuals actually observed to achieve that educational level. Although the treated are the group who receives most attention in the evaluation literature, for our purposes the ATT, representing the average payoff to individuals' own choices, is of particular relevance when compared to the average returns for other groups, most notably the 'non-treated'.

II. ***Marginal learners*: the non-treated**

If we focus on the 'non-treated', the corresponding parameter is the Average Effect of Treatment on the Non-Treated (ATNT), yielding the average return that those who have not achieved the educational level of interest would have enjoyed from undertaking that qualification. In terms of HE, for instance, the ATNT would thus tells us what the average return to HE would have been for non-graduates, had they undertaken HE. This is the parameter of interest if we wish to assess the impact of extending a currently voluntary educational attainment to the whole population, e.g. make it compulsory.

III. ***Marginal learners*: the eligibles**

If one is interested in the average return for all persons eligible to acquire a given educational qualification – irrespective of whether they have achieved it or not –, the parameter of interest is the Average Treatment Effect (ATE). This is the most relevant parameter if individuals were assigned to achieve the qualification *randomly* from the overall eligible population – as such this is probably not the most relevant parameter for evaluating individuals' educational choices. We will thus not focus on this parameter, and will not report our ATE estimates for the disaggregated analyses by income, social class or ability.[3]

IV. ***Marginal learners*: individuals with different attainment probabilities**

What characteristics should designate an individual as a marginal learner?
This is a question which can hardly be answered in a general context (we will discuss some specific definitions below), since it is the *interaction* of diverse individual and family background characteristics that come together in determining whether an individual achieves or not a given educational level. As a simple example, even though individuals from low- income families tend to be less likely to attain HE, ability tends to work in favour of educational attainment, so that low-income but high-ability individuals might in fact end up being just as likely to achieve HE as high-income but low-ability individuals.

One way around this issue is to consider a summary indicator of the likelihood that an individual with a given set of characteristics will undertake the educational qualification of interest. One such indicator is the 'propensity score', the probability of undertaking the qualification given observable characteristics. The propensity score for a given individual can thus be seen as an index summarising the *net* influence of the entire set of that individual's character-

[3] These are all available upon request.

istics on the likelihood that he or she will undertake the qualification of interest. Furthermore, the propensity score for a given individual can be easily estimated, and the predicted probability that the individual will take up the qualification can next be retrieved.

Appendix A outlines a possible way of exploiting the propensity score to define quite disaggregated groups of marginal learners, but the sample sizes in the BCS70 were just too small to allow a robust analysis along these lines.

Instead, we divided the 0-1 interval of the probability space into 4 equal-width bands of the educational attainment probability, and looked at the average returns to the corresponding educational level for groups who were increasingly more likely to achieve it, i.e. for:

- individuals who were the least likely to achieve the qualification (less than 25% chance);
- individuals with a medium-low attainment chance (25% to 50%);
- individuals with a medium-high attainment chance (over 50% but less than 75%);
- individuals who were the most likely to attain the qualification (over 75% chance).

V. ***Marginal learners*: defined a priori on a policy basis**

Our final way of proceeding is to define a number of alternative groups of marginal learners as those groups who are particularly targeted by policy in an attempt to raise their stay-on rates or HE participation.

Based on this 'policy definition', one could for instance be interested in the returns to education for individuals characterised by a given combination of ability and family income, or by a given socio-economic background. For example, individuals of high ability coming from a low-income family make up the potentially credit-constrained group, with cognitive and non-cognitive skills that could enable them to successfully continue past the minimum or to attain HE, but who because of their low family resources might have to forsake education for work. Similarly, individuals coming from a low-income family or from a low socio-economic background are the ones being generally targeted by educational policies and subsidies, such as the recently introduced EMA.

Compared to the analysis based on sub-dividing our entire eligible sample according to their probability of participating in education, this way of proceeding has the advantage of being based on a very clear definition of the group for whom we are calculating the returns, thereby offering the policy-maker clear targeting criteria based on one or two well-defined and observable characteristics. The drawback is that we no longer know the true probability that any one type will next undertake the qualification of interest.

The implementation of the estimation for the returns to 'policy defined' marginal learners is straightforward: the sample is subdivided into the subgroups of interest, for whom returns to educational qualifications are calculated separately. Note that one can then apply any of the four methods outlined above, and calculate returns for even 'more marginal' groups. In particular, one can estimate the ATT and ATNT separately. This would allow to recover, for instance, the average returns to HE for high-ability, low-income individuals who undertook HE (the ATT), as well what returns those high-ability, low-income individuals who did not go on to HE would have reaped had they undertaken it (ATNT).[4]

[4] Similarly one could estimate the returns for a given 'marginal' group, say low-income individuals, by

However, a severe limitation relates to sample sizes, which turned out to be particularly binding in the BCS70.[5]

To derive our careful and robust estimates – both of the ATT and ATNT – we thus had to focus on splitting the population of interest into two groups at a time:

- individuals from a high *versus* from a low socio-economic background,[6]
- individuals from a high *versus* from a low income family[7],
- individuals of high *versus* low ability.[8]

As a robustness check we also separately estimate the returns for individuals for whom we do not have income, parental class or ability information. These can only be loosely contrasted with those for the other groups, since the differences in estimates may arise partly from the selectivity of the non-responding groups and partly from omitted variable bias. In particular, attrition has been shown (see e.g. Fitzgerald, Gottschalk and Moffitt, 1998) to be highly selective and concentrated among individuals of lower socio-economic status, those with more variable earnings, and those who move more frequently. As to potential omitted variable bias, when interpreting the estimates it has to be kept in mind that for individuals with missing information on ability, income or parental social class we cannot obviously control for these factors as is the case for those individuals with complete information.

3. Methodology

In this section we briefly outline the methodology chosen and its underlying assumptions, mainly to be in a position to highlight some issues which are important for a correct interpretation of the empirical results (we refer to the Appendix A for an extensive and more formal discussion).

In order to estimate the average return to, say, HE for a specific group of individuals, one would ideally compare average wages for those individuals if they completed HE to average wages for those *same* individuals if they did not undertake HE. The obvious problem is that a given individual either undertakes HE or does not. Consequently, the average wage that graduates would have commanded had they not gone to HE, as well as the average wage that individuals who did not go on to HE would have commanded had they graduated are unobserved counterfactuals that need to be estimated through the use of appropriate methods and usually untestable assumptions.

---

educational attainment probability, such as the returns to HE for those individuals from a low social background whose other characteristics were such that they had a high probability of achieving HE.

[5] See the Appendix A for more details.

[6] We define an individual to be of 'high social class' if the father was a professional or skilled non-manual worker when the child was 16; of 'low social class' if the father was skilled manual, semi-skilled, unskilled, unemployed, a student, dead or otherwise absent. If information on father's social class at 16 was missing, it was replaced by the same information at age 10 whenever available, given the relatively high persistence of father's social class.

[7] Parental income at 16 is classified as 'high' if above the median and 'low' if below the median, where the median is calculated in relation to the population of interest (i.e. everyone or conditional on level 2).

[8] For sample size considerations, we require information on *at least one* test result at age 10. Individuals are defined to be of 'low' ability if they scored in the lowest two quintiles of any test they took at age 10, and of 'high' ability otherwise, provided they took at least one test at 10.

Simply comparing average wages for graduates to average wages of non-graduates might be misleading, since in general one might expect individuals going on to HE to do so on the basis of characteristics that also influence their wages. They might for instance be of higher ability than those who do not move on to HE, so that graduates would have earned, on average, more than non-graduates even if they had not acquired HE (the 'selection bias').

However if we have information detailed enough to capture all these relevant differences, we could select among non-graduates that sub-group which has the same overall characteristics of graduates. This is the idea behind the method of matching with its underlying conditional independence assumption (CIA). More in particular, if the evaluator has access to very rich background information capturing all those factors that jointly determine wages and HE achievement[9], matching methods can be used to construct a suitable comparison group.

Our empirical analyses are always performed separately by gender and our use of the BCS70 implicitly conditions on year of birth hence age. We explicitly control for a large number of characteristics of the individual, their family and their environment, which it is hoped should eliminate selection bias[10]: ethnicity, a variety of cognitive and non-cognitive measures at age 5, verbal, mathematical and general ability test scores at age 10, father's and mother's education, parental income and father's social class at 16, number of siblings and region.

As to the approach we use, it is important to highlight that matching (in contrast to standard parametric methods like OLS) is robust in the sense that it does not restrict at all the way in which the return to education may vary according to individual characteristics. In addition, the focus of matching methods is on the careful choice of an appropriate comparison group, and hence on the actual comparability of groups. This is achieved via two mechanisms, knowledge of which is important for the interpretation of the resulting estimates.

First, in order to find a suitable match, graduates need have a counterpart on the non-graduate population (the 'common support' condition). In fact, if there are graduates who are not comparable to anyone in the non-graduate group, they will need to be left unmatched. In such a situation, the estimated return has to be *redefined* as the mean return for those graduates who fall within the common support. This has the noteworthy implication that the return estimated over the common subset may no longer represent the ATT – the average return to HE for *all* graduates – but only the average return for that subset of graduates who fall within the common support.

The second mechanism is the re-weighing of the characteristics in the non-graduate group so as to realign them to those in the graduate group (once restricted to the common support). Note in fact that one can easily check how well matching has balanced the available observables between the two groups. If balancing cannot be achieved, the researcher needs to accept the fact that the treated and non-treated are simply too different in terms of the observables and that there simply is not enough information in the available data to achieve sufficiently close – and thus reliable – matches.

[9] Note that identification of the ATNT requires a more restrictive version of the CIA assumption than does the ATT, in particular one needs to rule out the possibility that individuals may select into education based on individual returns from it that are unobserved by the researcher (see the Appendix A for more details).

[10] Blundell, Dearden and Sianesi, B. (2005) find no evidence of remaining selection bias in the NCDS for the HE *versus* anything less decision once controlling for similar variables.

In order to summarise important details on matching quality for a given matching estimate and thus highlight its degree of reliability, in the Tables below we have formatted in italics those estimates that should be considered with care due in particular to either rejecting a test that all the characteristics are jointly well balanced in the two matched groups and/or to having lost 10% or more of the treated (or of the non-treated for ATNT) to the common support.

To conclude on the interpretational caveats, it is important to clearly highlight what our estimates represent (under the assumptions clarified above).

First, as mentioned in the introduction to this part, our estimates pertain to the *private* return to education accruing to the individual in the form of higher *wages*. This means that any potential externalities benefiting the economy at large, as well as any non-wage benefits accruing to the individual himself are not captured in our returns estimates. In addition, the individual wage returns to education we estimate are only one component in a full analysis of the private rates of return to education on wages, which balance individual costs against a *flow* of such returns over the working life. In focusing on *average* returns for specific subgroups we also do not take account of the riskiness of education returns, an important determinant of educational choices especially among less well-off families.

Finally, special caution is required when attempting a comparison of the returns by gender. Men are far more likely to be working in their early 30s than women, and men very rarely work part-time. As a result, when comparing returns for men and for women, we are comparing the returns for men working full-time to the returns for women, of whom almost a third works part-time. In addition, individuals might be likely to select into employment based in part on their returns from work. The effect of returns of education on the decision to work may be more important for women than men because staying out of the labour force is more of an option for women.

4. Results

We have considered the returns which accrue to different 'marginal' groups from two broad educational investments: staying on post-compulsory schooling (section 4.1) and attainment of an HE qualification (section 4.2) as compared to both at least level 2[11] and to anything less. Given that individuals in general need to have achieved at least some level 2 qualifications before moving on to HE, in our analysis of returns to HE we will mostly focus on the first type of comparison, highlighting potentially different results arising for the full comparison group where appropriate.

As motivated in the introduction, we view the more recent BCS70 as more policy relevant and thus focus our analysis on this cohort. The wage returns for the various 'marginal' subgroups of this cohort are all estimated at one point in time – in 1999/2000, when these individuals were aged 29-30. It has thus to be kept in mind that we are looking at one specific age group only, and not at returns over the life-cycle. There is in fact some evidence that returns to first degrees, and to a lesser extent A levels, may to be rising until individuals are aged in their early thirties, before stabilising thereafter (McIntosh, 2004).

[11] In constructing our level 2 qualification variable we include both academic and vocational qualifications. We also include individuals who have 5 or more GCSE or equivalent qualifications *at any grade*.

4.1 Staying on past compulsory schooling

We begin by looking at returns to staying on in school past age 16.

A first important finding is that among individuals who work, there is a considerable gender gap in staying on rates, with 54.7% of girls staying on as compared to 43.5% of boys. This disparity, though, does not primarily arise from differential selection in employment; in the entire (work and non-work sample) stay on rates are 40.9% for boys and 50% for girls.

For men, our OLS and matching estimates suggest that individuals who stay on in school earn around 11-12% more than if they had dropped out. The ATT and ATNT are nearly identical, indicating that if those who dropped out had instead stayed on, they would have received very similar returns to those who did stay on.

For women we find returns to staying on of around 18%, again coinciding for the staying-on group (ATT) and the dropping-out group (ATNT). Returns for women that are higher than those for men are a general finding in the education and training literature. As mentioned earlier, however, the returns for women are more difficult to interpret because women's educational choices and workforce participation decisions are more complicated than men's. Since women spend on average less time than men in the labour market and thus have less time on average to recoup their investment in education, it could be that those women who select into education are those with the highest returns from it. Additional problems in terms of selection into employment would further arise if those women with higher returns are more likely to be working.

We next break the sample into four bands based on their probability of staying on past 16 and investigate whether those who were very unlikely to stay in school, based on their observable characteristics, receive higher or lower returns than those who were very likely to stay past 16. Sample sizes make it very hard to obtain precise results, and in fact none of the differences in the within-gender estimates in the table are statistically significant at conventional levels. The differences in the point estimates of the returns among the four subgroups of men are found to be negligible, while the point estimate of the return for those women least likely to stay on is considerably smaller (though not statistically significantly so) than the others.

Moving to our policy characterisations of the marginal learner, we start by considering returns to staying on by father's social class. For men and women alike, returns are substantial for both the high and the low social class groups, with returns for the high social class group being somewhat (though not statistically significantly) larger than the gains for the lower social class group. In particular, men from a high socio-economic background have a 13-14% return from their staying on decision, compared to 8-11% for men from a low socio-economic background, while the corresponding figures for women are 17-19% compared to 15-16%. Of potential interest is the finding that men from a low socio-economic background who dropped out would have enjoyed the lowest return from staying on (8.4%) among these subgroups.[12]

[12] For completeness we also report the returns estimated for those individuals with missing information on parental social class (13% of the sample). Whilst almost as few of them stay on as individuals with reported low social class, their returns are one and a half times those for the low social class group. This is most likely to be due to an upward bias from not being able to control for social class and for other important characteristics these individuals have not reported (e.g. for 44% of this group information on ability at 10 is missing as well).

Another policy relevant group are those from a low-income family, in fact often targeted by educational subsidies and policies like the EMA. Parental income (which we measure at 16) is distinct from social class discussed above in that it should also capture transitory and more circumstantial influences on schooling decisions. Social class should by contrast represent more permanent and stable factors that our sample experienced in their youth.[13] It has to be warned that for half of the sample we do not observe parental income, making it unclear how generalisable the corresponding findings are. Appendix D reports on the results of a set of exploratory analyses performed to this regard.

Our OLS estimates for both men and women do not uncover large heterogeneity in returns by parental income, returns to staying on being respectively 11 and 13% for low- and high-income men, and 16 and 15% for low- and high-income women. The more flexible matching point estimates by contrast show some heterogeneity, both between the income divide and between treated and non-treated. Men from poorer families have an 11% return from their staying on decision, while men from richer ones have only a statistically insignificant 8% return. This 3 percentage points extra gain for the more disadvantaged group is however not statistically significant. In terms of the ATNT, the corresponding figures are around 13% for both income groups, highlighting how if men who dropped out had instead continued, they would have earned a higher return than those who did actually continue, and this quite irrespective of their family income. Again, however, the differences in the point estimates of the ATNT and the ATT for the two income groups do not pass a test of statistical significance.

As for women, from the point estimates it is those from richer backgrounds who are enjoying a higher return from their staying on decision (18% against 14%), but it is those from poorer backgrounds who would have enjoyed a higher return had they not dropped out (17% against 12%). None of these differences are however statistically significant.

For both genders, the returns to staying on are in any case substantial for the low-income group, and in particular for those who did not stay on. The analyses in Appendix D do however call for caution in interpreting the results for low-income males. Subject to this caveat, as well as to the caution required in extrapolating results to more recent periods (these estimates pertain to individuals who were making their staying-on choices in 1986), these findings would seem to imply that policies, such as the EMA, encouraging individuals from low-income families to stay on in school could be very beneficial to the target population.

We finally break the sample into two ability groups. Maybe surprisingly, neither OLS nor matching uncover any substantial heterogeneity in returns to staying on by individual ability, with men enjoying a 10% return and women a 15% return quite irrespective of their cognitive ability when they were 10. An interesting finding concerns men of lower ability; if those who dropped out at the minimum had instead continued, they would have enjoyed a mere 6.7% return – the lowest one compared to any sub-group we have considered, and in fact just over half of the return for those low-ability males who did decide to stay on, this difference passing our test of statistical significance. This finding indicates that a policy encouraging the less

[13] This conceptual distinction has prevented us from reducing the number of missing information as we did for social class. We felt it not appropriate to 'fill in' missing income information at 16 with the corresponding information at 10.

bright male pupils to stay on would not appear to be particularly efficient for enhancing private wage returns.[14]

[14] As to the 18% of the sample with missing ability information, the estimates display a huge upward biased, as one would expect when not conditioning on possibly the most important source of selection bias (for this group we do not have *any* ability test results at 10).

Table 1: Wage returns to staying on

	Males	Females
OLS	0.123 ***	0.181***
	(0.017)	(0.016)
ATT	0.111***	0.179***
	(0.057; 0.156)	(0.143; 0.222)
ATNT	0.107***	0.178***
	(0.071; 0.148)	(0.139; 0.215)
ATE	0.109***	0.179***
	(0.066; 0.142)	(0.145; 0.213)
ATNT-ATT	-0.004	-0.001
N	4150	3891

by Staying-on Probability

	Males		Females	
Band	Effect	95% conf.int.	Effect	95% conf.int.
0.00 to 0.25	0.117***	(0.058;0.206)	0.161**	(0.003;0.272)
0.25 to 0.50	0.144***	(0.077;0.196)	0.187***	(0.136;0.256)
0.50 to 0.75	0.138***	(0.053;0.206)	0.212***	(0.161;0.275)
0.75 to 1.00	*0.115**	(-0.005;0.229)	*0.192****	(0.089;0.290)

See Appendix B for the mean characteristics of individuals falling within each band and further statistics.

by Socio-economic Background

	Males			Females		
	Low	High	Missing	Low	High	Missing
OLS	0.101***	0.141***	0.174***	0.158***	0.174***	0.233***
	(0.021)	(0.027)	(0.039)	(0.021)	(0.027)	(0.046)
ATT	0.110***	0.126**	0.166***	0.151***	0.192***	0.251***
	(0.057;0.167)	(0.016;0.193)	(0.053;0.258)	(0.102;0.194)	(0.135;0.266)	(0.149;0.383)
ATNT	0.084***	0.141***	*0.161****	0.154***	0.169***	0.262***
	(0.039;0.135)	(0.075;0.208)	(0.045;0.301)	(0.107;0.205)	(0.109;0.228)	(0.157;0.404)
ATNT - ATT	-0.026	0.015		0.003	-0.024	
ATT(Low)-ATT(High)	-0.016			-0.041		
% stay-on	31	64	36	45	72	43
N (%)	2100 (51)	1501 (36)	549 (13)	1969 (51)	1456 (37)	466 (12)

Table 1: Wage returns to staying on (continued)

by Family Income

	Males			Females		
	Low	High	Missing	Low	High	Missing
OLS	0.112***	0.132***	0.159***	0.164***	0.154***	0.210***
	(0.030)	(0.032)	(0.021)	(0.029)	(0.032)	(0.021))
ATT	0.110***	0.080	0.177***	0.142***	0.179***	0.210***
	(0.040;0.175)	(-0.055;0.186)	(0.125;0.234)	(0.056;0.212)	(0.104;0.284)	(0.158;0.270)
ATNT	0.129***	0.125***	0.132***	0.172***	0.123***	0.207***
	(0.047;0.207)	(0.050;0.200)	(0.087;0.177)	(0.104;0.254)	(0.038;0.179)	(0.159;0.263)
ATNT-ATT	0.019	0.045		0.030	-0.056*	
ATT(Low)-ATT(High)	0.030			-0.037		
% stay-on	33	59	41	48	68	52
N (%)	1009 (24)	1037 (25)	2104 (51)	1020 (26)	1008 (26)	1863 (48)

by Ability

	Males			Females		
	Low	High	Missing	Low	High	Missing
OLS	0.104***	0.103***	0.261***	0.158***	0.156***	0.279***
	(0.023)	(0.025)	(0.035)	(0.023)	(0.025)	(0.036)
ATT	0.114***	0.093**	0.299***	0.155***	0.143***	*0.274* cs***
	(0.057;0.161)	(0.018;0.161)	(0.228;0.389)	(0.105;0.217)	(0.082;0.206)	(0.174;0.367)
ATNT	0.067***	0.108***	0.264***	0.173***	0.150***	0.261***
	(0.018;0.121)	(0.052;0.170)	(0.178;0.345)	(0.113;0.235)	(0.101;0.206)	(0.178;0.338)
ATNT-ATT	-0.047**	0.015		0.019	0.007	
ATT(Low)-ATT(High)	0.021			0.012		
% stay-on	25	62	45	40	68	58
N (%)	1736 (42)	1666 (40)	748 (18)	1618 (42)	1564 (40)	709 (18)

Notes to Table 1:

Data: BCS70, dependent variable is real log gross wages in 1999/2000.

Control variables: sex and age implicitly; ethnicity, quintile in a combined score of a number of ability tests taken at age 5, a composite measure of non-cognitive/social skills at age 5, separate quintiles in verbal, mathematical and general ability test scores at age 10, father's and mother's years of education, quintile of parental income at 16, father's social class at 16 (8 dummy variables), number of siblings and dummies for region of residence at 16.

Significance: *** at 1%, ** at 5% and * at 10%.

Socio-economic background: 'Low' if father was skilled manual, semi-skilled, unskilled, unemployed, student, dead or otherwise absent when child was 16; 'High' if father was professional or skilled non-manual worker. If information on father's social class at 16 was missing, it was replaced by the same information at age 10 whenever available.

Family income: 'Low' if below the median when child was 16; 'High' if above the median.

Ability: 'Low' if child scored in the lowest two quintiles of any test he/she took at age 10; 'High' otherwise, provided he/she took at least one test at 10.

4.2 Higher education

We now turn to our analysis of the returns to achieving higher education (HE) as compared to having stopped with qualifications at level 2 or 3.[15]

Our HE 'treatment' is defined broadly; it includes the attainment of any level 4 qualification, thus not just first degrees, but also other qualifications such as HNCs and diplomas of higher education. Our 'comparison treatment' is having obtained qualifications at level 2 or 3, which is often regarded as the pre-condition for moving on to HE. We thus consider the return in terms of individual wages from attaining any form of HE compared to stopping with at least a level 2 qualification. At the end of this section we briefly look into the results that arise when we estimate the returns to HE compared to anything less (the full results for this comparison are presented in Appendix C).

A first noteworthy finding is that the gender gap in educational attainment, which was particularly pronounced in staying-on rates, has completely vanished, with around half of the sample achieving some form of HE. Compared to the 10-11 percentage points gap displayed in staying-on rates by females compared to males, women are only marginally more likely (2-3 percentage points) than males to achieve HE conditional on level 2.[16] Nor is this due to our focusing on HE attainment conditional on level 2. If we just look at the unconditional HE participation in the population, again we find no gender gap.

Overall, we find substantial returns to HE conditional on at least level 2 for nearly every subsample we investigate, with the large majority of the point estimates being higher than the analogous returns to staying on.

For men, the OLS estimated return is 16%, the ATT 14% and the ATNT 16%. As was the case for staying on, returns for females are higher than the corresponding returns for men: female HE graduates enjoy a 22% return (ATT) from having undertaken HE, while female non-graduates with at least level 2 qualifications would have similarly enjoyed a 23% return had they gone on to HE (ATNT).

For both genders, we find that those with intermediate probabilities of achieving HE (between .25 and .75) experience statistically significantly higher returns from HE than those with either higher or lower HE attainment probabilities. If we define the marginal learners as those on the bubble between going on to HE and not, this result may indicate that policies directed at expanding HE places to accommodate this marginal group could help the beneficiaries achieve noticeably higher wages.

When we break the male sample into groups by socio-economic background, family income and ability, there are two important findings to note.

First, returns are considerably higher among the less advantaged men. In particular, low social class males have higher returns than high social class males (20% *versus* 9-11%), and low-income males have higher returns than high-income males (23-24% *versus* 9-12%).[17] When

[15] Note that we include both academic and vocational qualifications. We also include individuals who have 5 or more GCSE or equivalent qualifications *at any grade*.

[16] Conditional on at least level 2, 56.7% of working females have attained HE compared to 53.3% of working males. Irrespective of working status in 2000, the figures are 51% and 53.4% respectively.

[17] The exploratory analyses in Appendix D do not seem to imply that missing income poses serious problems for

we compare the ATT of the advantaged (high) and disadvantaged (low) groups, we find in fact that these differences between them are statistically significant.

To put this larger gain from HE for the disadvantaged males in perspective, it has to be noted that returns are measured as the *percentage* gain from the baseline 'counterfactual', and that low income and low social class males are starting from a much a lower baseline than their high income and high social class peers (see the Table below). Thus even though the proportionate gain from HE is higher for the disadvantaged male groups, it has to be borne in mind that the non-HE disadvantaged males earn much less (specifically, 20%), on average, than the non-HE advantaged males.

Real hourly gross wage of males and females without HE but at least level 2 qualifications, by ability, income and social class

	Males			Females		
	Ability	Income	Social class	Ability	Income	Social class
low	8.1	7.9	8.3	6.7	6.8	7.1
high	10.0	10.1	10.3	7.9	7.7	7.9

The second noteworthy result is that when we compare the returns to the treated and to the non-treated, we find that among low income men, the return to those attaining HE (ATT) is 5 percentage points higher than the return that those who did not attain HE would have enjoyed (ATNT), this extra gain being statistically significant. One possible explanation for this result is that the opportunity cost of HE is higher among the less well-off individuals and as a result their decision to attend HE is more well thought out and cognizant of the anticipated returns. Disadvantaged individuals thus seem to be already able to sort themselves into HE in an efficient way. Policies promoting HE among men from low income families may however still be justified on the basis of the substantial returns (around 19%) that these disadvantaged non-treated are foregoing.

Still comparing the returns to the treated and to the non-treated, we find that for low social class males, the ATT and the ATNT are roughly equivalent. By contrast, for both high social class and high income men, the ATNT exceeds the ATT, and statistically significantly so. In particular, the returns to HE would have been around 5 percentage points higher for those among them who did not undertake HE than for those who did undertake it. Subject to the usual caveat for the ATNT (see Appendix A), these results would seem to point to non-optimal HE decisions within the advantaged groups. One possible explanation is that for the well-off group, ex-post mistakes in HE participation decisions are less costly, so that such decisions need not always be based on expected returns and may in contrast often be driven by non-economic criteria.

Focusing on returns by ability group[18] for males, the patterns are the same as those we uncovered in terms of social class and income, though they are weaker in this dimension. In particular, we find slightly larger returns to HE for low-ability males than for high-ability males

the interpretation of the estimates for the low- and high-income groups of either gender.

[18] Note that the percent of the sample that is of low ability has dramatically fallen. While 42% of males as well as females were of low ability when in the previous section we included those who stopped below level 2, 33% of males as well as females are of low ability when we restrict the sample to level 2 or above. This is not surprising as low-ability individuals are less likely to continue on in school.

(17% *versus* 13% for ATT), though such difference is here not statistically significant. We also continue to find the ATT exceeding the ATNT for those of low ability, and the ATNT exceeding the ATT for those of high ability, although again these differences are quite small (2-3 percentage points) and not statistically significant.

When we turn to the results for females broken down by social class, family income and ability, we find that, in sharp contrast to males, women experience similar returns across categories. Advantaged and disadvantaged women on any of the three dimensions we consider experience surprisingly similar (and high) returns, all around 22-23%. For women, we also fail to find any set relationship between returns for the treated and for the non-treated. Instead, returns in all categories are basically equivalent for women who attain HE (ATT) and women who don't (ATNT).

These differing patterns for men and women may be due to selection into employment, whereby those women with the lowest returns, independent of income, social class or ability category are more likely to opt out of the labour market, and hence our estimation sample.

A final note relates to the extent of missing information. Although missing values are an enduring feature of the data even for this more educated set of individuals, the biases are considerably lower than in the previous section. This illustrates how restricting our attention to individuals with at least level 2 qualifications already controls for several important determinants of educational investments as well as wages. Similarly, the analyses in Appendix D indicate that in terms of HE outcomes, both males and females from low- as well as from high-income families do not systematically differ in terms of unobservables from their missing-income counterparts.

The patterns of findings we have discussed so far as to the returns to HE compared to at least level 2 which accrue to variously defined groups of 'marginal' entrants are overall very similar to those in terms of the returns to HE compared to anything below HE. The complete findings are presented in Appendix C; in what follows we just highlight the general patterns and main lessons from this further analysis.

In particular, as already mentioned, we find no gender gap in HE attainment in the population, though of course attainment rates are now considerably lower given our focus on a less educated group.[19]

A second obvious point is that when we compare HE recipients to the whole population of non-HE recipients, we expect to find larger returns than when we compared HE to those who had attained at least level 2 qualifications, again due to our comparison group being less educated. In fact, most of the point estimates for returns, falling between 15% and 30%, are higher than the analogous returns conditional on at least level 2.

[19] Compared to over 50% HE attainment rates for individuals with at least level 2, around one third of men and women achieve some form of HE. The exact figures are 36% for working men and 38% for working women. Considering the entire sample irrespective of working status in 2000, males and females still have virtually the same HE achievement rates, 33% and 32% respectively.

Our matching estimates indicate that both the treated and the non-treated enjoy substantial returns of 18% and 21% for males and 25% and 27% for females, with the slightly higher point estimates of the ATNT being not statistically significantly different from the ATT.

Once the full sample is broken down into groups by social class, family income and ability, we find a very similar pattern of results to the one which emerged when we compared HE to at least level 2.

In particular, for males we still find that the returns to HE are considerably and statistically significantly higher among the less advantaged, that is they are higher for low social class men, for low income men and for low ability men, than for their more advantaged peers, all these differences being statistically significant.

In parallel with the results for HE *versus* at least level 2, the ATT is significantly higher than the ATNT for low-income males. Our point estimates suggest the opposite result among high social class and high income men; in contrast to our earlier results, although these differences are quite large (4-5 percentage points), they are not statistically significant.

We also uncover the same relationship between ATT and ATNT within both ability groups; this time however, it is in this comparison that such differences are statistically significant. More specifically, low-ability non-HE males would have had 5 percentage points lower returns than the low-ability males who did attain HE, which is consistent with individually optimal choices. However, among high-ability men, estimated returns to HE tend to be higher for those who do not achieve HE than for those who do. This indicates that sorting is not efficient and that policies encouraging bright individuals to continue on to HE could yield substantial returns to the target population.[20]

Finally, results for women are very much in line with those obtained when restricting our comparison group to individuals holding at least level 2 qualifications, namely that returns for women in the low and high social class, by income and by ability are broadly similar, as are the ATT and ATNT in each of the categories. Again these results may in part be due to selection into employment among women.

[20] This finding of sub-optimal HE decisions among the high-ability group is quite difficult to rationalise; one possibility – in addition to a violation of the ATNT assumption – relates to informational failures or high discount rates (in the form of either 'impatience' or credit constraints) among the non-treated high-ability individuals, who miss out on their high returns to HE since they are content to accept jobs which even without higher qualifications are sufficiently high-paying for them in view of their higher ability. It is under this interpretation that there would seem to be scope for policy to encourage or assist bright individuals in pursuing HE.

Table 2: Wage returns to HE compared to level 2 or above

	Males	Females
OLS	0.157***	0.231***
	(0.019)	(0.018)
ATT	0.142***	0.220***
	(0.100; 0.180)	(0.178; 0.258)
ATNT	0.159***	0.234***
	(0.113; 0.200)	(0.198; 0.273)
ATE	0.150***	0.226***
	(0.112; 0.187)	(0.186; 0.259)
ATNT-ATT	0.017	0.015
N	2811	2571

by HE-attainment Probability

	Males		Females	
Band	Effect	95% conf.int.	Effect	95% conf.int.
0.00 to 0.25	*0.016*	(-0.152;0.072)	0.111	(-0.182;0.262)
0.25 to 0.50	0.204***	(0.150;0.280)	0.297***	(0.255;0.365)
0.50 to 0.75	0.169***	(0.110;0.241)	0.223***	(0.113;0.267)
0.75 to 1.00	*0.039*	(-0.160;0.123)	*0.152***	(0.014;0.237)

See Appendix B for the mean characteristics of individuals falling within each band and further statistics.

by Socio-economic Background

	Males			Females		
	Low	High	Missing	Low	High	Missing
OLS	0.196***	0.111***	0.140***	0.225***	0.232***	0.200***
	(0.024)	(0.028)	(0.044)	(0.025)	(0.027)	(0.061)
ATT	0.198***	0.090**	*0.144***	0.231***	0.231***	*0.203***
	(0.148;0.254)	(0.011;0.157)	(0.040;0.240)	(0.180;0.285)	(0.160;0.298)	(0.008;0.358)
ATNT	0.191***	0.138***	*0.130***	0.225***	0.233***	*0.209****
	(0.137;0.250)	(0.086;0.207)	(0.008;0.219)	(0.160;0.274)	(0.172;0.294)	(0.055;0.426)
ATNT-ATT	-0.008	0.048**		-0.006	0.002	
ATT(Low)-ATT(High)	0.108***			0.000		
% HE	44	63	52	48	67	51
N (%)	1257 (45)	1214 (43)	340 (12)	1144 (45)	1163 (45)	264 (10)

Table 2: Wage returns to HE compared to level 2 or above – (continued)

by Family Income

	Males			Females		
	Low	High	Missing	Low	High	Missing
OLS	0.227***	0.118***	0.146***	0.230***	0.220***	0.246***
	(0.035)	(0.034)	(0.024)	(0.034)	(0.033)	(0.026)
ATT	0.244***	0.086	0.152***	0.218***	0.201***	0.230***
	(0.181;0.350)	(-0.024;0.140)	(0.096;0.205)	(0.109;0.287)	(0.125;0.282)	(0.165;0.297)
ATNT	0.192***	0.139***	0.144***	0.226***	0.223***	0.257***
	(0.092;0.257)	(0.078;0.222)	(0.091;0.188)	(0.159;0.308)	(0.145;0.295)	(0.206;0.308)
ATNT-ATT	-0.053*	0.053**		0.008	0.023	
ATT(Low)-ATT(High)	0.158***			0.017		
% HE	47	61	52	49	65	55
N (%)	630 (22)	823 (29)	1358 (48)	612 (24)	768 (30)	1191 (46)

by Ability

	Males			Females		
	Low	High	Missing	Low	High	Missing
OLS	0.162***	0.147***	0.165***	0.234***	0.219***	0.265***
	(0.028)	(0.025)	(0.042)	(0.030)	(0.025)	(0.041)
ATT	0.167***	0.126***	*0.210***	0.231***	0.224***	0.263***
	(0.092;0.226)	(0.066;0.188)	(0.139;0.313)	(0.145;0.297)	(0.169;0.284)	(0.170;0.357)
ATNT	0.145***	0.156***	0.172***	0.233***	0.225***	0.265***
	(0.088;0.201)	(0.096;0.208)	(0.086;0.301)	(0.175;0.304)	(0.174;0.283)	(0.153;0.355)
ATNT-ATT	-0.022	0.030		0.002	0.001	
ATT(Low)-ATT(High)	0.041			0.006		
% HE	40	60	59	44	64	61
N (%)	919 (33)	1376 (49)	516 (18)	847 (33)	1246 (48)	478 (19)

Notes to Table 2:

Data: BCS70, dependent variable is real log gross wages in 1999/2000.

Control variables: sex and age implicitly; ethnicity, quintile in a combined score of a number of ability tests taken at age 5, a composite measure of non-cognitive/social skills at age 5, separate quintiles in verbal, mathematical and general ability test scores at age 10, father's and mother's years of education, quintile of parental income at 16, father's social class at 16 (8 dummy variables), number of siblings and dummies for region of residence at 16.

Significance: *** at 1%, ** at 5% and * at 10%.

Socio-economic background: 'Low' if father was skilled manual, semi-skilled, unskilled, unemployed, student, dead or otherwise absent when child was 16; 'High' if father was professional or skilled non-manual worker. If information on father's social class at 16 was missing, it was replaced by the same information at age 10 whenever available.

Family income: 'Low' if below the median when child was 16; 'High' if above the median. Median calculated conditional on at least level 2.

Ability: 'Low' if child scored in the lowest two quintiles of any test he/she took at age 10; 'High' otherwise, provided he/she took at least one test at 10.

5. Conclusions

In this concluding section we summarise our main findings for the BCS70 cohort in terms of the individual wage returns to staying on and to HE enjoyed by variously defined 'marginal' groups when aged 29/30.

Our general finding is that there are substantial returns to both staying on and to HE for all subgroups of the population, though we have uncovered some variation in returns, especially for men. Returns are invariably found to be higher for women than for men – as we have argued throughout, this ranking needs to be interpreted with caution.

As to staying on past compulsory schooling:

- Our estimates of the returns to staying on are around 11% for men and 18% for women.

- For either males of females, we have not uncovered statistically significant heterogeneity in returns by actual staying-on decision, staying-on probability, social class, family income or ability.

- The lowest returns overall (6-8%) are those that male drop-outs of either low-ability or low social class would have enjoyed had they stayed on.

- For both mean and women, individuals from a low-income family that drop out would have enjoyed substantial returns from staying on (around 13% for men and 17% for women). This result should be interpreted with care given the patterns of missing income information. If we take the additional caution required when extrapolating results for the mid 80s to a more recent period, this finding is indicative that individuals from poorer backgrounds who do not stay in school could considerably benefit were they encouraged to do so.

As to the attainment of any form of higher education conditional on having achieved at least level 2 qualifications:

- We find sizeable average wage returns to HE relative to holding at least a level 2 qualification, estimates being around 15% for men and 22% for women. (Returns to HE compared to anything less are obviously larger, around 20% for men and 25% for women).

  In fact, nearly all of our estimates of returns to HE are higher than the returns to staying on for the same subgroups.

- In terms of returns by attainment probability, we find that it is those men and women who are 'indifferent' (i.e. in the 25-50% probability range) between undertaking HE or remaining at level 2 or 3 who experience the highest returns from moving on to HE.

- For men, we find that the returns to HE are substantially higher for the more disadvantaged groups.

  In particular, low socio-economic class men enjoy significantly higher returns than high social class men; men from a low-income family have significantly higher returns than those from a more well-off family; and low-ability men have slightly higher returns to HE than high-ability men (this latter difference being statistically significant only for HE compared to anything less). Note that we are estimating percentage returns; these proportionate gains – so much higher for disadvantaged men – have to be viewed in a context where disadvantaged individuals without HE earn much less, on average, than advantaged individuals without HE.

- Among low-income men, we find higher returns to HE for those who did achieve HE than for those who stopped at level 2 or 3, which might be taken as an indication of informed, efficient sorting into HE.

- By contrast, for high-income males, as well as for men from high social class, the average HE return the non-treated would have enjoyed exceeds the average return enjoyed by the treated, this difference being statistically significant. One possible explanation relates to an expanded role of non-economic factors in the decision-making of the well-off groups.

- For the low social-class male sample, the average returns to HE for the treated and for the non-treated are roughly equivalent.

- In sharp contrast to men, for women we find very similar returns to HE (compared both to at least level 2 and to anything less) across ability, income and social class groups, as well as for each treated and non-treated subgroups. This may be due to factors concerning selection into employment.

6. References

Blundell, R., Dearden, L. and Sianesi, B. (2005), "Evaluating the Impact of Education on Earnings in the UK: Models, Methods and Results from the NCDS", *Journal of the Royal Statistical Society*, Series A, forthcoming.

Card, D. (2001), "Estimating the Returns to Schooling: Progress on Some Persistent Econometric Problems", *Econometrica*, 69, 1127–60.

Dearden, L., McGrahanan, L. and Sianesi, B. (2004), "The Role of Credit Constraints in Educational Choices: Evidence from the NCDS and BCS70", forthcoming Centre for Economics of Education Discussion Paper.

Fitzgerald, J., Gottschalk, P. and Moffitt, R. (1998), "An Analysis of the Impact of Sample Attrition on the Second Generation of Respondents in the Michigan Panel Study of Income Dynamics", *The Journal of Human Resources* 33, no. 2: 300-344

Heckman, J.J., LaLonde, R. and Smith, J. (1999), "The Economics and Econometrics of Active Labor Market Programs", in O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, vol. 3, Amsterdam: Elsevier-North Holland.

McIntosh, S. (2004), "Further Analysis of the Returns to Academic and Vocational Qualifications", Centre for Economics of Education Discussion Paper No. 35

Appendix A – Methodological approach to estimate returns to the marginal learner

The evaluation problem

The problem of measuring the returns to an educational qualification falls neatly into the evaluation framework, where the aim is to measure the causal effect of the educational 'treatment' on the outcome of interest *Y*, which for us is individual wages (see, for example, Card, 2001, and Heckman, LaLonde and Smith, 1999).

Letting Y_{1i} be the wage of individual *i* if s/he achieved the qualification and Y_{0i} the wage if *i* were not to achieve the qualification, the causal effect (or return) for individual *i* of achieving the qualification is $Y_{1i}-Y_{0i} \equiv \beta_i$. However, since no individual can be in two different educational states at the same time, either Y_{1i} or Y_{0i} is missing for each *i*, which makes it impossible to ever observe the individual return β_i. Our more modest though still challenging aim is to identify the *average* returns in some population of interest *T*, that is $E(\beta_i \mid i \in T)$

An important decision to be made at this stage is whether to assume homogeneous or heterogeneous individual returns. In fact if one were willing to assume homogeneous returns across individuals ($Y_{1i}-Y_{0i} = \beta$ for all individuals *i*), the returns to the sub-populations of marginal entrants, however defined, would by construction all coincide.

However one would in general not expect all students to benefit from a given educational investment in exactly the same way: there will most likely be heterogeneity across individuals in the returns to a given qualification. Once the return is allowed to vary across individuals, the immediate question concerns the choice of the population *T* for which to estimate average returns.

To formalise the following discussion, let $D_i \in \{0,1\}$ be a binary indicator of achievement of the qualification of interest by individual *i*. Note then that the actually observed wage for individual *i* is given by $Y_i = Y_{0i} + D_i \cdot (Y_{1i}-Y_{0i})$, i.e. the no-qualification wage Y_{0i} plus, if the individual achieves the qualification ($D_i=1$), the return β_i from it. Finally, let X_i be a set of observed characteristics (not affected by education) of individual *i*.

Identification of our parameters of interest

The Average Effect of Treatment on the Treated is the average return for those individuals who have chosen to undertake the educational qualification of interest:

$$\text{ATT} \equiv E(Y_{1i} - Y_{0i} \mid D_i=1) \equiv E(\beta_i \mid D_i=1) = E(Y_{1i} \mid D_i=1) - E(Y_{0i} \mid D_i=1)$$

For individuals acquiring the qualification, say HE to fix ideas, we do observe Y_1, their post-HE wages, so that the average observed wage of those individuals having achieved HE is an unbiased estimate of the first component of the ATT. We do not however observe graduates' Y_0, the wage they would command had they not undertaken HE. To overcome this fundamental identification problem, the average counterfactual $E(Y_0 \mid D=1)$ needs to be somehow constructed on the basis of some usually untestable identifying assumptions that justify the use of the observed pairs (Y_1, D=1), (Y_0, D=0).

One such assumption and the one we used to derive our estimates is the conditional independence assumption (CIA), requiring the educational decision to be independent of the no-education out-

come conditional on a set of observed characteristics X, in symbols $Y_0 \perp D \mid X$. The CIA thus requires the evaluator to observe all those characteristics that affect *both* the educational choice *and* potential outcomes.

Under the CIA, matching offers a way to construct a suitable comparison group. Matching involves explicitly selecting and pairing to each HE graduate a non-graduate with the 'same' X, or, more generally, attaching appropriate weights to the observations in the non-graduate group, so as to realign the distribution of X in the non-graduate group to the one in the group of graduates.

For the matching procedure to have empirical content, the Common Support condition is also required: $P(D=1 \mid X)<1$. This condition thus prevents X from being a perfect predictor of HE achievement, guaranteeing that all graduates have a counterpart on the non-graduate population for the set of X values over which we seek to make a comparison. Depending on the sample in use, this can be quite a strong requirement (for example, when the education level under consideration is directed to a well-specified group). If there are regions where the support of X does not overlap for the graduate and non-graduate groups, matching has to be performed over the common support region, and the estimated return has then to be *redefined* as the mean return for those graduates who fall within the common support. It is important to note that if the return to HE differs across graduate individuals, restricting to the common subset may actually change the parameter being estimated. In other words, the estimated return may no longer represent the average return to HE for *all* graduates (the ATT), but only the average return for that subset of graduates who fall within the common support.

Note that although both matching and simple OLS regression rely on the potentially strong CIA assumption, matching is not subject to several potential misspecification biases for the ATT compared with standard parametric methods like OLS. In particular, OLS may suffer from mis-specification bias for the non-education outcome equation; it may use this imposed functional form to extrapolate outside the common support, if need be, and thus de facto compare observationally different individuals; and in general it does not identify the ATT in the presence of heterogeneous effects.

We now turn to the Average Effect of Treatment on the Non-Treated, i.e. the average return among those who have not achieved the educational level of interest:

$$\text{ATNT} \equiv E(Y_{1i} - Y_{0i} \mid D_i=0) \equiv E(\beta_i \mid D_i=0) = E(Y_{1i} \mid D_i=0) - E(Y_{0i} \mid D_i=0)$$

Identification of the ATNT requires identification of the counterfactual = $E(Y_{1i} \mid D_i=0)$, that is how much would have non-graduates earned, on average, had they achieved HE.

To identify such a counterfactual the CIA is required to hold in terms of the HE outcome Y_1 (i.e. independence of the educational decision D from wages if education is undertaken, given observed characteristics X), while the Common Support requirement becomes $P(D=1 \mid X)>0$.[21]

It is important to stress that the identification of the ATNT requires a more restrictive CIA assumption than does the ATT, in particular one needs to rule out the possibility that individuals may select into education based on (from us) unobserved individual returns from it. More specifically, if given the information in the available X, there is still some individual idiosyncratic gain which is unobserved by the econometrician but known in advance (or predictable with accuracy) by the individual when making his or her educational choices, then it would seem sensible to assume that choices

[21] For the ATNT, the imposition of the common support thus means that the effect has to be redefined as the mean return those low-education individuals who fall within the common support would have enjoyed had they further invested in education.

will, in part at least, reflect the return of that choice. If this is the case – and assuming that it is those with the higher idiosyncratic gains who select into HE – the matching estimate of the ATNT will be *upward* biased.

How likely is this condition to be violated? Such a discussion should first take into account the kind of information which is captured by X; the answer would then depend on what we are willing to believe that individuals know, at the time of their educational decision, about their future *realised* idiosyncratic (i.e. over and above what is captured by X) component of the return.

The Average Treatment Effect is the expected return for an individual randomly selected from the overall eligible population:

$$\text{ATE} \equiv E(Y_{1i} - Y_{0i}) \equiv E(\beta_i)$$

Since the ATE is just a weighted average of the ATT and ATNT[22], its identification hinges on the correct identification of *both* the ATT and the ATNT, which imposes more stringent behavioural assumptions in term of the CIA as well as data requirements in terms of the Common Support Conditions.

We now consider the estimation of the returns to a qualification by attainment probability. A summary indicator of the likelihood that an individual with a given set of characteristics will undertake the educational qualification of interest is the 'propensity score', the conditional probability of undertaking the qualification given observable characteristics X:

$$p_i \equiv P(D_i=1 \mid X_i)$$

The propensity score for a given individual can easily be estimated, e.g. using a probit or logit model. One can then retrieve the predicted probability $\hat{p}_i$ that individual i with characteristics X_i will undertake the qualification of interest.

One could then directly determine those individuals who had propensity scores closest to the minimum associated with attainment in the high-education group, and estimate the average returns for them (area A in Figure 1). Alternatively, one could identify those individuals who had propensity scores closest to the maximum observed in the low-education group, and again estimate the average returns for them (area B in Figure 1).

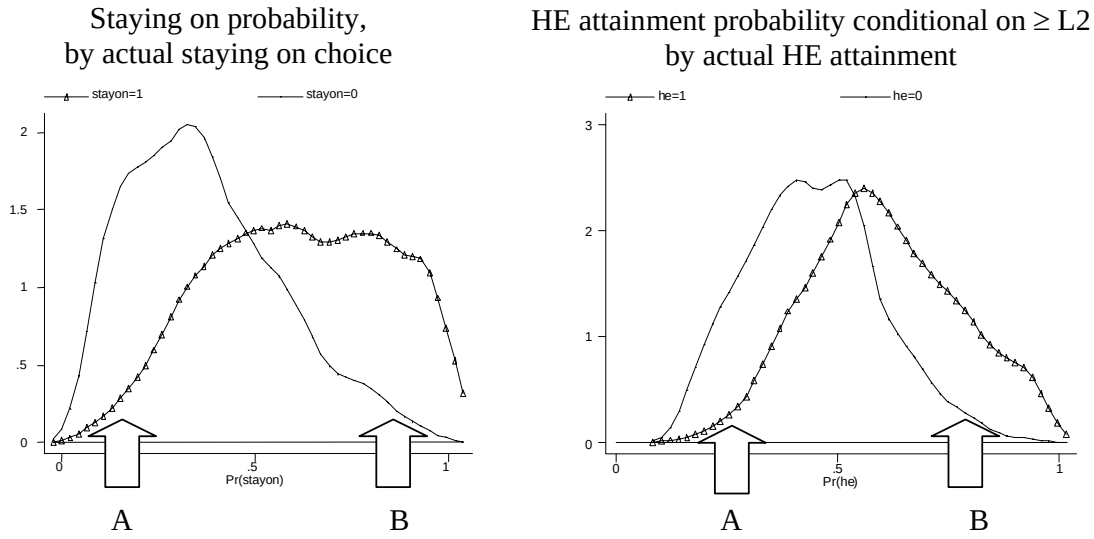
Unfortunately, the sample sizes in the cohort data are simply too small to yield a disaggregated analysis of this type which is sufficiently robust. In particular, note that the number of high-education individuals falling around A, as well as of low-education individuals falling around B would be by construction very small.

We thus pursued a different strategy. We estimated the average returns to a given educational qualification for groups of individuals whose attainment probability fell within a given interval $[m, n]$, $0<m<n<1$, by the difference between the mean wage of the treated falling within that band and the mean wage of the non-treated falling within that band:

$$E(Y_{1i}-Y_{0i} \mid \hat{p}_i \in [m, n]) \equiv E(\beta_i \mid \hat{p}_i \in [m, n]) = E(Y_i \mid D_i=1,\ \hat{p}_i \in [m, n]) - E(Y_i \mid D_i=0,\ \hat{p}_i \in [m, n])$$

[22] $E(Y_1 - Y_0) = E(Y_1 - Y_0 \mid D=1)P(D=1) + E(Y_1 - Y_0 \mid D=0)\cdot(1-P(D=1))$

Figure 1(b) Density distribution of the propensity score in the BCS70 data



A couple of comments should be helpful as to the interpretation of the results from this type of analysis.

First, although the single characteristics of individuals defined by a given value (or band) of the propensity score will be different, it is obviously possible to look at the mean characteristics of these individuals and get a flavour of the type of individuals they are. It could thus be possible to draw an average 'profile' for individuals whom the policymaker may wish to target based on their attainment / stay-on probability. We present a brief summary of such characteristics in Appendix D.

The second clarification concerns the interpretation of the returns within propensity score bands as being estimates of the ATT, ATNT or ATE for individuals falling within that band. The matching framework allows the returns to arbitrarily vary according to characteristics X (not so the standard OLS model, which assumes/imposes the return to be the same, irrespective of X). However, for individuals with the *same* set of characteristics X, matching assumes homogenous returns. Due to an important property of the propensity score, the returns turn out to be homogenous also for individuals with the same value of the propensity score, which as we have seen can be viewed as a summary of the full set of X. Hence for the results by propensity score p it makes no sense to distinguish between ATT, ATNT or ATE, since by construction they all coincide for that value of p: $ATT(p) = ATNT(p) = ATE(p) \equiv \beta_p$. However, we discussed above that the conditions to identify the ATT are less stringent than those required to identify the ATNT or the ATE; in particular, individuals are allowed to select into education based on their unobserved returns from it. It is thus more robust to interpret the estimates by propensity score as ATT estimates, i.e. as estimates of the average return to, say, HE attainment for HE graduates whose attainment probability fell in a given interval.

Turning finally to the implementation of our policy-driven definition of the marginal learner, the sample of the eligible population is subdivided into the subgroups of interest, for whom returns to educational qualifications are calculated separately. For each policy subgroup one can then use matching methods to recover estimates of the returns for even 'more marginal' groups. In particular, for a given policy sub-group, say low-income individuals, one can separately estimate the ATT, the ATNT and the returns by educational attainment probability.

As we mentioned in the main text, the binding limitation turned out to relate to the small sample sizes available in the BCS70. If those individuals who are working (with non-missing wages) and have non-missing information both on ability and parental income at 16 are separated by gender and further into four high/low ability/income cells, the resulting cell sizes are simply too small to allow for robust non-parametric methods such as matching (the smallest cell has 211 observations, the largest 554). OLS could be able to yield more precise estimates even for small cell sizes due to the structure it imposes in terms of linearity, homogeneous returns and disregard of common support. We feel however that these imposed restrictions could seriously undermine the reliability of the results, while additional serious concerns would remain as to the representativeness and generability of these estimates, given potential selection into response.

Appendix B – Returns by propensity score band

a) Staying on

		Males				Females			
From	To	$P>\chi$ (1)	#Stay on	#No Stay on	%Stay on	$P>\chi$ (1)	#Stay on	#No Stay on	%Stay on
0.00	0.25	0.057	169	898	15.8	0.323	49	258	16.0
0.25	0.50	0.256	375	689	35.2	0.830	434	650	40.0
0.50	0.75	0.545	468	290	61.7	0.080	707	451	61.1
0.75	1.00	0.011	522	72	87.9	0.000	660	112	85.5

(1) p-value of joint test of balancing of the *X* variables between the two groups

Average individual characteristics by propensity score band

	Males				Females			
	[0.00;0.25]	[0.25;0.50]	[0.50;0.75]	[0.75;1.00]	[0.00;0.25]	[0.25;0.50]	[0.50;0.75]	[0.75;1.00]
Non-white	0.000	0.007	0.049	0.096	0.000	0.006	0.041	0.073
Pa yrs education	3.327	4.773	6.522	10.519	3.094	4.375	6.103	10.933
Ma yrs education	3.388	4.892	6.590	10.359	3.153	4.483	6.288	10.975
Pa education miss	0.663	0.539	0.405	0.194	0.687	0.568	0.419	0.130
Ma education miss	0.662	0.530	0.393	0.178	0.684	0.557	0.396	0.108
Pa Professional I	0.000	0.007	0.032	0.194	0.000	0.003	0.010	0.166
Pa Intermediate II	0.022	0.076	0.201	0.409	0.007	0.041	0.157	0.439
Pa Skilled non-manual III	0.013	0.059	0.083	0.086	0.007	0.042	0.074	0.073
Pa Skilled manual III	0.201	0.237	0.204	0.098	0.173	0.231	0.270	0.149
Pa Semi-skilled IV	0.055	0.046	0.042	0.022	0.091	0.058	0.046	0.013
Pa Unskilled manual V	0.019	0.005	0.004	0.003	0.029	0.023	0.009	0.003
Pa Unknown SC	0.067	0.097	0.103	0.061	0.062	0.088	0.107	0.066
No pa/unemp/student/dead	0.022	0.046	0.049	0.039	0.042	0.054	0.052	0.039
Pa SC/empl missing	0.600	0.428	0.282	0.089	0.590	0.460	0.274	0.053
Income – bottom quintile	0.246	0.166	0.102	0.054	0.215	0.199	0.159	0.085
Income – 2nd quintile	0.126	0.096	0.090	0.039	0.153	0.136	0.096	0.061
Income – 3rd quintile	0.062	0.081	0.098	0.096	0.062	0.065	0.098	0.079
Income – 4th quintile	0.035	0.098	0.139	0.231	0.042	0.058	0.127	0.193
Income – top quintile	0.020	0.048	0.082	0.239	0.010	0.019	0.054	0.247
Income missing	0.512	0.511	0.491	0.342	0.518	0.522	0.465	0.334
No siblings	0.110	0.104	0.096	0.084	0.085	0.102	0.104	0.098

1 sibling	0.208	0.293	0.369	0.487	0.173	0.254	0.385	0.479
2 siblings	0.184	0.190	0.178	0.237	0.192	0.192	0.200	0.238
≥3 siblings	0.145	0.109	0.116	0.098	0.173	0.140	0.115	0.115
Sibling information missing	0.353	0.304	0.240	0.094	0.378	0.312	0.196	0.069
BAS test – bottom quintile	0.305	0.079	0.045	0.013	0.635	0.149	0.038	0.008
BAS test – 2nd quintile	0.255	0.153	0.074	0.027	0.287	0.306	0.105	0.040
BAS test – 3rd quintile	0.245	0.203	0.121	0.054	0.059	0.240	0.212	0.100
BAS test – 4th quintile	0.100	0.212	0.266	0.209	0.003	0.138	0.272	0.240
BAS test – top quintile	0.013	0.112	0.352	0.604	0.003	0.030	0.193	0.472
BAS test score missing	0.082	0.241	0.141	0.093	0.013	0.137	0.180	0.141
reading test – bottom quintile	0.410	0.054	0.034	0.007	0.678	0.081	0.018	0.009
reading test – 2nd quintile	0.326	0.181	0.067	0.030	0.228	0.315	0.107	0.026
reading test – 3rd quintile	0.157	0.268	0.175	0.089	0.081	0.324	0.161	0.104
reading test – 4th quintile	0.047	0.212	0.297	0.204	0.010	0.142	0.289	0.223
reading test – top quintile	0.005	0.060	0.305	0.599	0.000	0.022	0.258	0.505
reading test score missing	0.056	0.225	0.121	0.071	0.003	0.115	0.166	0.133
maths test – bottom quintile	0.335	0.061	0.026	0.008	0.521	0.154	0.044	0.018
maths test – 2nd quintile	0.354	0.140	0.037	0.013	0.322	0.300	0.115	0.057
maths test – 3rd quintile	0.164	0.227	0.090	0.044	0.075	0.227	0.185	0.109
maths test – 4th quintile	0.060	0.208	0.274	0.148	0.003	0.113	0.250	0.246
maths test – top quintile	0.003	0.070	0.380	0.641	0.003	0.026	0.177	0.399
maths test score missing	0.084	0.293	0.193	0.145	0.075	0.180	0.230	0.171
ability at 5 – bottom quintile	0.297	0.117	0.055	0.019	0.397	0.168	0.076	0.039
ability at 5 – 2nd quintile	0.236	0.137	0.119	0.061	0.202	0.200	0.139	0.096
ability at 5 – 3rd quintile	0.143	0.180	0.177	0.145	0.107	0.174	0.177	0.145
ability at 5 – 4th quintile	0.083	0.172	0.189	0.173	0.068	0.155	0.208	0.196
ability at 5 – top quintile	0.034	0.141	0.224	0.369	0.010	0.064	0.179	0.338
ability at 5 – missing	0.206	0.254	0.236	0.234	0.215	0.239	0.221	0.187
Social skills at 5	0.333	-0.029	-0.191	-0.298	0.337	-0.095	-0.236	-0.334
regional health authority=1	0.087	0.047	0.037	0.025	0.029	0.054	0.064	0.071
regional health authority=2	0.060	0.053	0.067	0.074	0.101	0.065	0.053	0.045
regional health authority=3	0.097	0.071	0.055	0.051	0.114	0.094	0.065	0.051
regional health authority=4	0.055	0.039	0.026	0.035	0.059	0.030	0.031	0.025
regional health authority=5	0.029	0.033	0.038	0.077	0.020	0.028	0.044	0.066
regional health authority=6	0.034	0.029	0.045	0.034	0.016	0.032	0.041	0.041
regional health authority=7	0.035	0.047	0.055	0.054	0.023	0.035	0.049	0.073
regional health authority=8	0.005	0.034	0.036	0.086	0.003	0.008	0.029	0.091
regional health authority=9	0.030	0.034	0.038	0.039	0.039	0.032	0.046	0.028
regional health authority=10	0.016	0.022	0.042	0.047	0.033	0.037	0.029	0.043
regional health authority=11	0.036	0.041	0.041	0.057	0.081	0.069	0.054	0.053
regional health authority=12	0.085	0.086	0.086	0.079	0.078	0.101	0.081	0.061
regional health authority=13	0.035	0.035	0.041	0.037	0.039	0.043	0.046	0.047
regional health authority=14	0.066	0.051	0.050	0.040	0.046	0.055	0.065	0.048
regional health authority=15	0.040	0.056	0.066	0.077	0.003	0.018	0.051	0.102
regional health authority=16	0.040	0.083	0.100	0.128	0.026	0.051	0.099	0.133
Region missing	0.251	0.240	0.175	0.059	0.290	0.246	0.154	0.022

b) HE conditional on at least level 2 qualifications

		Males				Females			
From	To	$P>\chi$ (1)	#Stay on	#No Stay on	%Stay on	$P>\chi$ (1)	#Stay on	#No Stay on	%Stay on
0.00	0.25	0.000	23	126	15.4	0.950	14	31	31.1
0.25	0.50	0.505	362	542	40.0	0.192	299	465	39.1
0.50	0.75	0.847	602	388	60.8	0.537	655	403	61.9
0.75	1.00	0.002	277	57	82.9	0.003	296	63	82.5

(1) p-value of joint test of balancing of the *X* variables between the two groups

Average individual characteristics by propensity score band

	Males				Females			
	[0.00;0.25]	[0.25;0.50]	[0.50;0.75]	[0.75;1.00]	[0.00;0.25]	[0.25;0.50]	[0.50;0.75]	[0.75;1.00]
Non-white	0.000	0.003	0.028	0.123	0.000	0.012	0.044	0.072
Pa yrs education	4.497	4.372	6.524	12.284	3.844	5.251	6.890	12.298
Ma yrs education	4.403	4.470	6.623	11.961	4.200	5.488	6.946	12.092
Pa education miss	0.550	0.577	0.406	0.114	0.622	0.484	0.372	0.114
Ma education miss	0.564	0.569	0.396	0.096	0.578	0.463	0.357	0.097
Pa Professional I	0.000	0.000	0.011	0.359	0.000	0.009	0.036	0.248
Pa Intermediate II	0.027	0.069	0.232	0.395	0.111	0.097	0.224	0.465
Pa Skilled non-manual III	0.020	0.048	0.085	0.045	0.022	0.060	0.075	0.042
Pa Skilled manual III	0.322	0.241	0.188	0.039	0.222	0.277	0.203	0.081
Pa Semi-skilled IV	0.027	0.039	0.043	0.048	0.022	0.047	0.038	0.011
Pa Unskilled manual V	0.027	0.011	0.000	0.000	0.000	0.010	0.009	0.011
Pa Unknown SC	0.094	0.108	0.092	0.024	0.022	0.058	0.107	0.058
No pa/unemp/student/dead	0.020	0.038	0.044	0.024	0.000	0.039	0.066	0.045
Pa SC/empl missing	0.463	0.447	0.304	0.066	0.600	0.402	0.242	0.039
Income – bottom quintile	0.255	0.184	0.101	0.054	0.311	0.211	0.121	0.047
Income – 2nd quintile	0.141	0.115	0.095	0.042	0.067	0.120	0.098	0.067
Income – 3rd quintile	0.154	0.101	0.081	0.048	0.133	0.088	0.084	0.045
Income – 4th quintile	0.040	0.067	0.167	0.222	0.000	0.079	0.142	0.201
Income – top quintile	0.013	0.039	0.087	0.329	0.000	0.017	0.095	0.345
Income missing	0.396	0.494	0.470	0.305	0.489	0.486	0.460	0.295
No siblings	0.128	0.142	0.069	0.051	0.200	0.126	0.096	0.072
1 sibling	0.282	0.260	0.387	0.518	0.111	0.285	0.405	0.479
2 siblings	0.181	0.187	0.206	0.237	0.244	0.241	0.179	0.223
≥3 siblings	0.181	0.139	0.090	0.093	0.022	0.085	0.128	0.184
Sibling information missing	0.228	0.272	0.248	0.102	0.422	0.263	0.193	0.042
BAS test – bottom quintile	0.369	0.113	0.026	0.012	0.667	0.169	0.021	0.003
BAS test – 2nd quintile	0.289	0.207	0.049	0.027	0.156	0.274	0.080	0.031
BAS test – 3rd quintile	0.208	0.263	0.093	0.054	0.178	0.304	0.125	0.031
BAS test – 4th quintile	0.047	0.191	0.249	0.216	0.000	0.152	0.286	0.189
BAS test – top quintile	0.020	0.123	0.388	0.536	0.000	0.033	0.283	0.613
BAS test score missing	0.067	0.103	0.194	0.156	0.000	0.069	0.205	0.134
reading test – bottom quintile	0.658	0.103	0.020	0.003	0.556	0.123	0.016	0.003
reading test – 2nd quintile	0.201	0.272	0.070	0.042	0.311	0.245	0.061	0.017
reading test – 3rd quintile	0.114	0.270	0.140	0.084	0.111	0.339	0.113	0.050
reading test – 4th quintile	0.013	0.195	0.244	0.204	0.022	0.207	0.250	0.209
reading test – top quintile	0.007	0.080	0.349	0.524	0.000	0.034	0.362	0.596
reading test score missing	0.007	0.081	0.176	0.144	0.000	0.052	0.197	0.125
maths test – bottom quintile	0.470	0.110	0.014	0.015	0.956	0.154	0.017	0.008

maths test – 2nd quintile	0.362	0.227	0.040	0.021	0.044	0.284	0.083	0.036
maths test – 3rd quintile	0.054	0.219	0.103	0.060	0.000	0.217	0.164	0.081
maths test – 4th quintile	0.007	0.212	0.226	0.108	0.000	0.174	0.252	0.198
maths test – top quintile	0.000	0.060	0.406	0.599	0.000	0.034	0.240	0.515
maths test score missing	0.107	0.173	0.210	0.198	0.000	0.136	0.244	0.162
ability at 5 – bottom quintile	0.490	0.147	0.028	0.018	0.378	0.144	0.071	0.025
ability at 5 – 2nd quintile	0.134	0.148	0.160	0.069	0.311	0.225	0.105	0.053
ability at 5 – 3rd quintile	0.168	0.210	0.137	0.108	0.111	0.183	0.164	0.095
ability at 5 – 4th quintile	0.054	0.143	0.199	0.159	0.044	0.168	0.216	0.189
ability at 5 – top quintile	0.000	0.091	0.264	0.413	0.000	0.075	0.218	0.446
ability at 5 – missing	0.154	0.261	0.212	0.234	0.156	0.205	0.225	0.192
Social skills at 5	0.574	0.085	-0.244	-0.289	-0.130	-0.149	-0.244	-0.311
regional health authority=1	0.020	0.042	0.063	0.051	0.000	0.056	0.070	0.084
regional health authority=2	0.020	0.051	0.070	0.093	0.022	0.025	0.054	0.114
regional health authority=3	0.161	0.079	0.051	0.048	0.111	0.072	0.077	0.081
regional health authority=4	0.074	0.038	0.031	0.045	0.022	0.026	0.032	0.025
regional health authority=5	0.094	0.042	0.022	0.039	0.000	0.018	0.046	0.070
regional health authority=6	0.047	0.028	0.038	0.018	0.022	0.018	0.043	0.056
regional health authority=7	0.087	0.043	0.045	0.036	0.022	0.037	0.051	0.064
regional health authority=8	0.000	0.022	0.037	0.105	0.000	0.037	0.040	0.045
regional health authority=9	0.020	0.027	0.041	0.048	0.044	0.030	0.040	0.033
regional health authority=10	0.007	0.015	0.032	0.069	0.067	0.045	0.027	0.028
regional health authority=11	0.054	0.060	0.033	0.039	0.044	0.081	0.050	0.053
regional health authority=12	0.060	0.076	0.097	0.102	0.067	0.111	0.079	0.058
regional health authority=13	0.034	0.040	0.041	0.033	0.067	0.052	0.050	0.036
regional health authority=14	0.027	0.041	0.067	0.078	0.022	0.050	0.058	0.070
regional health authority=15	0.074	0.062	0.062	0.054	0.022	0.035	0.057	0.067
regional health authority=16	0.047	0.127	0.080	0.096	0.111	0.105	0.085	0.106
Region missing	0.174	0.208	0.189	0.048	0.356	0.202	0.141	0.011

Appendix C – Wage returns to HE compared to anything less

	Males	Females
OLS	0.206*** (0.016)	0.265*** (0.016)
ATT	0.182*** (0.140; 0.219)	0.247*** (0.203; 0.279)
ATNT	0.207*** (0.161; 0.248)	0.268*** (0.225; 0.301)
ATE	0.198*** (0.158; 0.231)	0.260*** (0.221; 0.288)
ATNT-ATT	0.025	0.021
N	4150	3891

by HE-attainment Probability

	Males		Females	
Band	Effect	95% conf.int.	Effect	95% conf.int.
0.00 to 0.25	*0.245***	(0.177;0.322)	0.274***	(0.192;0.334)
0.25 to 0.50	*0.257***	(0.215;0.342)	0.297***	(0.249;0.359)
0.50 to 0.75	0.181***	(0.093;0.247)	0.245***	(0.170;0.312)
0.75 to 1.00	0.059	(-0.104;0.221)	*0.201***	(0.029;0.363)

by Socio-economic Background

	Males			Females		
	Low	High	Missing	Low	High	Missing
OLS	0.239*** (0.021)	0.160*** (0.025)	0.240*** (0.037)	0.263*** (0.022)	0.264*** (0.024)	0.307*** (0.048)
ATT	0.241*** (0.193;0.287)	0.131*** (0.054;0.190)	0.243*** (0.133;0.316)	0.268*** (0.223;0.313)	0.257*** (0.199;0.319)	0.276*** (0.151;0.378)
ATNT	0.212*** (0.150;0.262)	0.175*** (0.112;0.234)	*0.207*** (0.125;0.285)	0.265*** (0.209;0.313)	0.268*** (0.221;0.329)	0.282*** (0.155;0.391)
ATNT-ATT	-0.029	0.044		-0.003	0.011	
ATT(Low)-ATT(High)	0.110***			0.011		
% HE	27	51	32	28	53	29
N (%)	2100 (51)	1501 (36)	549 (13)	1969 (51)	1456 (37)	466 (12)

Wage returns to HE compared to anything less (continued)

by Family Income

	Males			Females		
	Low	High	Missing	Low	High	Missing
OLS	0.256***	0.153***	0.217***	0.262***	0.266***	0.290***
	(0.029)	(0.030)	(0.021)	(0.032)	(0.029)	(0.022)
ATT	0.271***	0.120***	0.213***	0.250***	0.237***	0.269***
	(0.213;0.345)	(0.035;0.188)	(0.164;0.260)	(0.162;0.315)	(0.174;0.294)	(0.204;0.321)
ATNT	0.216***	0.165***	0.205***	0.258***	0.271***	0.316***
	(0.117;0.275)	(0.093;0.242)	(0.142;0.247)	(0.187;0.345)	(0.199;0.346)	(0.279;0.368)
ATNT-ATT	-0.055*	0.045		0.008	0.034	
ATT(Low)-ATT(High)	0.150***			0.013		
% HE	29	48	33	30	50	35
N (%)	1009 (24)	1037 (25)	2104 (51)	1020 (26)	1008 (26)	1863 (48)

by Ability

	Males			Females		
	Low	High	Missing	Low	High	Missing
OLS	0.217***	0.188***	0.242***	0.260***	0.255***	0.340***
	(0.024)	(0.023)	(0.35)	(0.026)	(0.023)	(0.034)
ATT	0.232***	0.158***	0.241***	0.265***	0.244***	0.293***
	(0.181;0.297)	(0.109;0.210)	(0.150;0.321)	(0.213;0.326)	(0.185;0.296)	(0.205;0.349)
ATNT	0.181***	0.214***	0.221***	0.263***	0.268***	0.348***
	(0.120;0.239)	(0.160;0.260)	(0.123;0.286)	(0.195;0.320)	(0.223;0.322)	(0248;0.432)
ATNT-ATT	-0.051**	0.056**		-0.001	0.024	
ATT(Low)-ATT(High)	0.074**			0.021		
% HE	21	50	40	23	51	41
N (%)	1736 (42)	1666 (40)	748 (18)	1618 (42)	1564 (40)	709 (18)

Notes to Appendix C:

Data: BCS70, dependent variable is real log gross wages in 1999/2000.

Control variables: sex and age implicitly; ethnicity, quintile in a combined score of a number of ability tests taken at age 5, a composite measure of non-cognitive/social skills at age 5, separate quintiles in verbal, mathematical and general ability test scores at age 10, father's and mother's years of education, quintile of parental income at 16, father's social class at 16 (8 dummy variables), number of siblings and dummies for region of residence at 16.

Significance: *** at 1%, ** at 5% and * at 10%.

Socio-economic background: 'Low' if father was skilled manual, semi-skilled, unskilled, unemployed, student, dead or otherwise absent when child was 16; 'High' if father was professional or skilled non-manual worker. If information on father's social class at 16 was missing, it was replaced by the same information at age 10 whenever available.

Family income: 'Low' if below the median when child was 16; 'High' if above the median.

Ability: 'Low' if child scored in the lowest two quintiles of any test he/she took at age 10; 'High' otherwise, provided he/she took at least one test at 10.

Appendix D – Individuals with missing parental income at 16

In this Appendix we look at individuals with and without parental income information at 16, trying to assess how 'different' these two groups are.

A brief summary of the main finding of these analysis is as follows:

- Missing-income individuals are more likely to have other information missing, though particularly problematic is social class rather than ability. Overall, however, for 9% of this sample we have information on neither income, ability nor social class.
- Though obviously not two random samples, missing-income individuals are overall not too dissimilar from individuals with non-missing income once their diverse pattern of other missing information is taken into account.
- Missing-income individuals seem to either have a predicted parental income distribution which is remarkably similar to the observed income distribution of the group with observed income, or they are slightly more likely to come from the lowest quintile. These conclusions equally apply to the treated and non-treated, in terms of both staying on and HE.
- Missing-income individuals do not seem to significantly differ from individuals with non-missing income in terms of unobservables affecting their educational outcomes. There is one important exception however: low-income men do seem to differ from men with missing income in terms of some unobservable which affects their staying-on probability, though not their HE attainment probability. Staying-on results for low-income males need thus to be interpreted with care.

We start by considering their potentially diverse patterns of missing information on other important individual characteristics; this should inform us of the actual extent to which we can control for such information when deriving our matching estimates.

Table D1 summarises the patterns of missing ability and social class information for individuals with and without income information. As expected, individuals with missing income are more likely to have other information missing. This is relatively much more the case for social class than for our ability measure. The proportion who misses ability but not social class is in fact basically the same (around 12%) for individuals with missing income and those with non-missing income. Note however that for a much larger proportion of individuals with missing income we do not observe either ability or social class; the estimates in the text thus need to be taken with special care, since for over 9% of the missing-income sample we cannot control for ability, social class and parental income.

Individuals with and without parental income information are obviously not random subsamples. In testing how different these two groups are in terms of their other observable characteristics, however, the overall picture emerging from Table D2 is that individuals with missing information are overall not too dissimilar from individuals with non-missing income information once their diverse pattern of other missing information is taken into account. In particular, they do not seem to significantly differ in many dimensions, this being the case especially for females. And when there are significant differences, these seem mostly to arise from the missing information in that dimension.

Table D1: Patterns of missing ability and social class information for our full BCS70 sample, individuals with non-missing income and individuals with missing income at 16

	Full sample	Individuals with non-missing income	Individuals with missing income
Missing ability	18.2%	14.1%	22.2%
Missing soc class	12.6%	6.6%	18.8%
Missing ability only	12.6%	12.3%	12.9%
Missing soc class only	7.1%	4.7%	9.5%
Missing both ability and soc class	5.5%	1.9%	9.3%
Missing neither ability nor soc class	74.8%	81.2%	68.3%
	100%	100%	100%

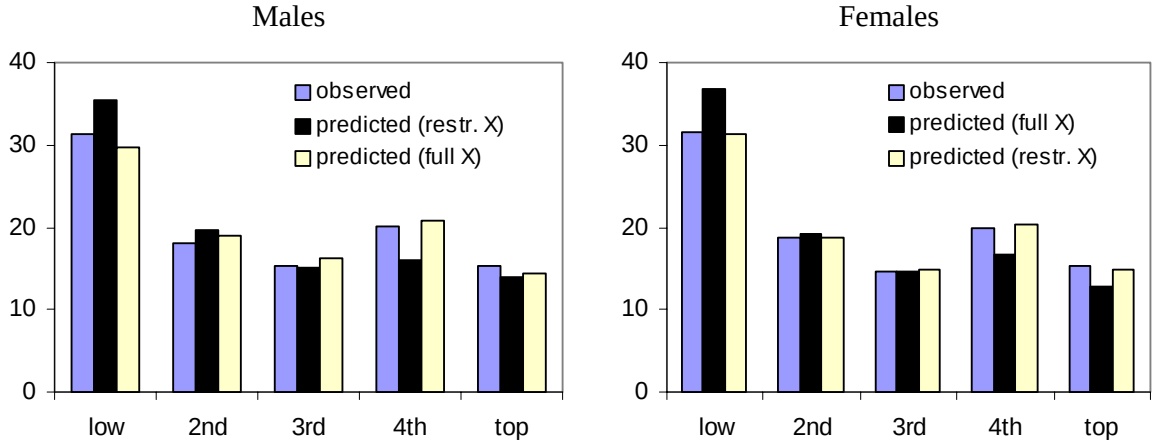
Table D2: Significance of the differences between individuals with missing information and individuals with non-missing information (p-value of joint significance tests, based on robust standard errors)

	Full set of controls		Restricted set of controls (1)	
	Males	Females	Males	Females
Race	0.2621	0.1106	0.8623	0.3155
All ability tests	0.0104**	0.8882	0.0000***	0.1520
– not missing only (2)	0.1013	0.8902	0.0140**	0.9228
– Tests at 10	0.0105**	0.8464	0.0000***	0.7828
– Tests at 5	0.9265	0.7807	0.6323	0.3435
– Social skills	0.3694	0.8491	0.2178	0.9715
Family background	0.0000***	0.0000***	0.0000***	0.0000***
– Siblings	0.2580	0.6138	–	–
– Father's social class	0.0000***	0.0000***	0.0001***	0.0000***
not missing only (2)	0.0335**	0.0217**	0.0094***	0.1700
– Parental education	0.0000***	0.0000***	0.0000***	0.0000***
not missing only (2)	0.4056	0.0670*	0.3517	0.0731*
Region	0.0054***	0.2750	–	–
Observations	2903	2943	4150	3891

Notes:
Significance: *** at 1%, ** at 5% and * at 10%.
Marginal effects and significance of all the individual variables are available upon request.
(1) Missing region or missing sibling information predict missing income perfectly. If we include these regressors we lose 43% of the male sample and 32% of the female sample.
(2) Not missing only: testing joint significance of the corresponding group of variables excluding the missing indicator.

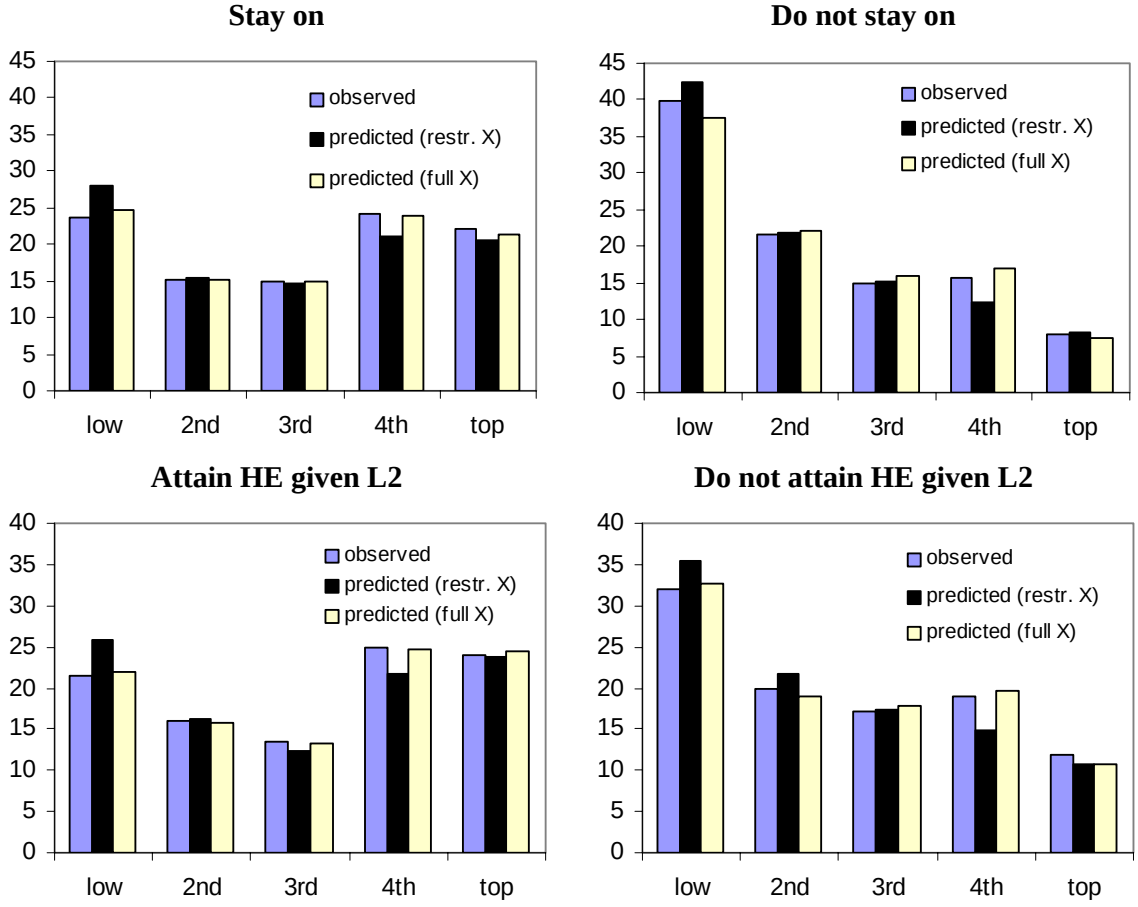
For the missing-income group, parental income is an unobservable we cannot control for. What would the distribution of parental income look like for the missing-income group? Would these individuals mainly come from the lower part of the distribution? If we assume that non-response in income is driven by our rich observed characteristics, we can in fact answer this question by looking at the income distribution of those individuals with non-missing income who most closely resemble our missing-income individuals in terms of all other observables. This is accomplished via matching, which as it turns out balances all our observables extremely well between the missing-income group and its matched non-missing-income subgroup.

Figure D1: Individuals with non-missing income: distribution of parental income (quintiles) for the full sample (observed) and for the subsample matched to the missing-income group (predicted – i.e. predicted for the missing-income group), separately by gender



Note: Restricted *X*: we exclude region and siblings information; when using the full set of *X* variables we lose 43% of the male sample and 32% of the female sample.

Figure D2: Individuals with non-missing income: distribution of parental income (quintiles) for the full sample (observed) and for the subsamples matched to the missing-income group (predicted – i.e. predicted for the missing-income group), separately by educational outcomes



Note: We have lumped gender to save space; the patterns are very much the same.

Figure D1 shows that overall, the predicted income distribution for the missing-income group is not too different from the observed income distribution of the group with income information, this being especially the case when matching on the full set of control variables. When we do not control for region and sibling composition, individuals with missing information seem more likely to come from the lowest quintile of the parental income distribution and less likely to be in the 4th (and top) income quintile.

Interestingly, Figure D2 confirms that both these patterns keep holding for both the treated and non-treated in terms of our two main educational outcomes.

Up to now we have focused on potential differences in *observable* characteristics between individuals with and without missing income. Some information as to how different these two groups are in terms of unobservables can be gleaned by looking at whether, once we control for our observables, the two groups display similar staying on and HE attainment rates. Conditioning on our observables, are the two groups just as likely to achieve the educational level to which we measure the returns in the paper, or do they still differ? In particular, if after controlling for observed characteristics the two subgroups significantly and systematically differ in terms of educational outcomes, they must differ in terms of some unobservable.

Table D3 shows that in terms of the raw differences in educational outcomes, individuals with missing income are less likely to either stay on or to achieve HE. The notable exception is that once we condition on having achieved at least a level 2 qualification, they no longer display a significantly different probability of attaining HE. This is in line with our general finding that conditioning on level 2 already selects a much more comparable set of people.

Once we control for our set of *X* variables, the non-significant small difference in HE rates conditional on level 2 decreases further, while the difference now disappears also unconditional on level 2: individuals with missing income information become just as likely to attain HE than individuals with non-missing income. For females, this result also holds in terms of staying-on rates, so that female with missing and non-missing income but the same *X*'s are just as likely to stay on. By contrast, men with missing income are significantly more likely to stay on than those men with non-missing income but with their same other characteristics *X*.[23] Men with missing income thus seem to differ from men with non-missing income in terms of some unobservable which affects their staying-on probability, though not their HE attainment probability.

These conclusions hold irrespective of whether we use the full sample and do not control for region and siblings or whether we perform the analysis on the reduced sample we are left with once we do control for region and siblings.

We finally investigate how different low-income individuals are from missing-income individuals in terms of unobservables, as well as how high-income individuals differ from missing-income individuals in terms of their unobservables.

This is accomplished along the same lines as above, and the results are presented in Tables D4 and D5.

[23] In fact, once we compare men with missing income to those men with non-missing income but with their same other characteristics *X*, men with missing income become 4-5 percentage points *more* likely to stay, i.e. selection on our observables *X* has reversed the situation in terms of staying on rates.

Table D3: Differences in educational outcomes between individuals with missing income information and individuals without missing income information: Raw differences and adjusted estimates from matching controlling for restricted set of *X* / full set of *X*

Educational outcome		Males	Females
Staying on	raw	-0.050***	-0.0627***
	adjusted	0.0491** / 0.0458**	0.0200 / 0.0054
HE given L2	raw	-0.0295	-0.0288
	adjusted	0.0078 / -0.0035	-0.0024 / 0.0153
HE	raw	-0.0544***	-0.0423***
	adjusted	0.0148 / 0.0058	0.0151 / 0.0236

Notes: Significance: *** at 1%, ** at 5% and * at 10%.
The degree of achieved balancing of the covariates (not shown) is unusually high across all comparisons.
Restricted *X*: we exclude region and siblings information; when using the full set of *X* variables we lose 43% of the male sample and 32% of the female sample.

Table D4: Differences in educational outcomes of low-income individuals compared to missing-income individuals: Raw differences and adjusted estimates from matching controlling for restricted set of *X* / full set of *X*

Educational outcome		Males	Females
Staying on	raw	-0.083***	-0.033*
	adjusted	-0.052*** / -0.071***	-0.033 / -0.029
HE given L2	raw	-0.046*	-0.058**
	adjusted	0.021 / 0.021	-0.043 / -0.042
HE	raw	-0.040**	-0.057***
	adjusted	0.004 / -0.006	-0.041* / -0.042*

Notes: Significance: *** at 1%, ** at 5% and * at 10%.
The degree of achieved balancing of the covariates (not shown) is unusually high across all comparisons.
Restricted *X*: we exclude region and siblings information; when using the full set of *X* variables we lose 43% of the male sample and 32% of the female sample.

Table D5: Differences in educational outcomes of high-income individuals compared to missing-income individuals: Raw differences and adjusted estimates from matching controlling for restricted set of *X* / full set of *X*

Educational outcome		Males	Females
Staying on	raw	0.179***	0.160***
	adjusted	0.030 / -0.001	-0.033* / -0.025
HE given L2	raw	0.087***	0.098***
	adjusted	0.006 / 0.002	0.022 / -0.015
HE	raw	0.146***	0.142***
	adjusted	0.033* / 0.015	0.021 / -0.017

Notes: Significance: *** at 1%, ** at 5% and * at 10%.
The degree of achieved balancing of the covariates (not shown) is unusually high across all comparisons.
Restricted *X*: we exclude region and siblings information; when using the full set of *X* variables we lose 43% of the male sample and 32% of the female sample.

The findings in terms of raw differences are as expected, with low-income individuals being less likely to achieve our educational outcomes than individuals with missing income (by 4 to 8 percentage points) and with high-income individuals are considerably more likely to attain further and higher education than individuals with missing income (by 9 to 18 percentage points).

Once we condition on our set of observables, these differences mostly disappear, with the notable exception of staying-on rates for males from low-income families.[24] For this group and this outcome, controlling for our set of observables is not enough: missing-income males with the same other observables characteristics as low-income males are still significantly more likely to stay on (but interestingly, as was the result for the full sample, they are not more likely to attain HE).

[24] Possibly also of HE *versus* anything less for low-income females.

CENTRE FOR THE ECONOMICS OF EDUCATION
Recent Discussion Papers

44	F. Galindo-Rueda O. Marcenaro-Gutierrez A. Vignoles	The Widening Socio-Economic Gap in UK Higher Education
43	S. Machin S. McNally	The Literacy Hour
42	S. Gibbons S. Machin	Paying for Primary Schools: Supply Constraints, School Popularity or Congestion?
41	J. Blanden P.Gregg	Family Income and Educational Attainment: A Review of Approaches and Evidence for Britain
40	A. Chevalier	Parental Education and Child's Education: A Natural Experiment
39	A. Jenkins	Women, Lifelong Learning and Employment
38	A. Jenkins R. Levacic	Evaluating the Effectiveness of Specialist Schools
37	A. Jenkins A. Wolf	Regional Variations in Adult Learning and Vocational Training: Evidence from NCDS and WERS 98
36	L. Feinstein F. Galindo-Rueda A. Vignoles	The Labour Market Impact of Adult Education and Training: A Cohort Analysis
35	S. McIntosh	Further Analysis of the Returns to Academic and Vocational Qualifications
34	J.S. Pischke	The Impact of the School Year on Student Performance and Earnings: Evidence from the German Short School Year
33	A. Chevalier G. Conlon	Does it Pay to Attend a Prestigious University?

32	F. Galindo-Rueda A. Vignoles	Class Ridden or Meritocratic? An Economic Analysis of Recent Changes in Britain
31	F. Galindo-Rueda	Employer Learning and Schooling-Related Statistical Discrimination in Britain
30	R. Dur C. Teulings	Are Education Subsides an Efficient Redistributive Device?
	A. Chevalier G. Conlon F. Galindo-Rueda S. McNally	Research Report to the Department for Education & Skills: The Returns to Higher Education Teaching
29	A. Jenkins A. Wolf	The growth of psychometric testing for selection: why has test use increased, will growth continue and what does this mean for education?
28	Arnaud Chevalier Gauthier Lanot	Monotonicity and the Roy Model

To order a discussion paper, please contact the
Publications Unit
Centre for Economic Performance
Tel 020 7955 7673 Fax 020 7955 7595