

**The Widening Socio-economic Gap in UK
Higher Education**

Fernando Galindo-Rueda

Oscar Marcenaro-Gutierrez

Anna Vignoles



CENTRE FOR THE
ECONOMICS OF
EDUCATION

December 2004

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

© Fernando Galindo-Rueda, Oscar Marcenaro Gutierrez and Anna Vignoles,
submitted September 2004

ISBN 07530 1738 5

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

In 1998, up front tuition fees were introduced for degree courses in the UK. Although poorer students are exempt from such fees, or at least pay lower amounts, many commentators predicted that fees would cause a fall off in the number of students entering Higher Education (HE) *and* that this would affect poorer students disproportionately. On the other hand, many economists have argued that, quite apart from the efficiency and equity arguments in favour of the introduction of tuition fees for HE (see for example Barr and Crawford (1998); Dolton, Greenaway and Vignoles (1997); Greenaway and Haynes (2003)), the huge wage gains from a degree, combined with a relatively low tuition fee, would be unlikely to put students off going to university, regardless of their socio-economic background.

This paper provides up to date empirical evidence on the socio-economic gap in HE participation, for the period spanning the introduction of tuition fees. We assess whether the gap has widened and ask whether the socio-economic gap emerges on entry into university or much earlier in the education system. We do this in two ways. Firstly we consider the likelihood of going to university if you come from a poor neighbourhood and analyse changes in this likelihood over time. Secondly, we use more detailed individual level data to model the determinants of HE participation, focusing on changes in the relationship between family background and HE participation over time.

Children from all socio-economic backgrounds are considerably more likely to go to university in 2001, as compared to 1994. In fact the growth in HE participation amongst poorer students has been remarkably high, mainly because they were starting from such a low base. Nonetheless our results suggest that poorer neighbourhoods (postcodes) saw a less rapid growth in the number of young people enrolled in HE as compared to richer neighbourhoods, particularly in the early and mid 1990s. The strength of the relationship between neighbourhood income levels and HE enrolment grew most rapidly in the early part of the period, rather than after the introduction of tuition fees. This would seem to imply that any income-driven inequality in HE is part of a longer-term trend, perhaps related to the gradual reduction in student support in

HE and the big expansion of the university sector that occurred in the early 1990s

Our detailed individual level analysis suggests that in 1996, i.e. before tuition fees, there was certainly substantial social class educational inequality in HE but that it occurred largely as a result of inequalities earlier in the education system. By 2000 however, one can observe social class effects on HE participation, even after conditioning for the number of GCSEs and A levels that an individual has. This seems to suggest a widening of the social class gap in higher education itself in the period after the introduction of tuition fees. However, in models that include finer measures of educational achievement, the social class effects become smaller and insignificant. We conclude that much of the impact from social class on university attendance actually occurs well before entry into HE. Of course just because we observe inequality in attainment at earlier ages, does not mean it is not related to problems in HE. Students may look forward and anticipate barriers to participation in HE and make less effort in school as a result. This is an area that requires further research.

The Widening Socio-economic Gap in UK Higher Education

Fernando Galindo-Rueda
Oscar Marcenaro-Gutierrez
Anna Vignoles

1	Introduction	1
2	Higher Education Policy in the UK	3
3	Trends in Higher Education Participation	4
4	Methodology and Data	5
	4.1 Higher education statistics agency data	5
	4.2 Youth cohort study data	7
5	Neighbourhood Income Levels and Changes in HE Participation	8
6	Micro-Analysis of the Determinants of HE Participation	12
	6.1 GCSE comparator group	14
	6.2 Changes over time	17
7	Conclusions	17
	References	19
	Figs	23
	Tables	26
	Appendices	31

Acknowledgements

The authors are grateful for extensive comments from the Department for Education and Skills and the Higher Education Funding Council.

Fernando Galindo-Rueda is a Research Officer at the Centre for Economic Performance, London School of Economics and Centre for the Economics of Education. Oscar Marcenaro Gutierrez is a Research Officer at the Centre for Economic Performance, London School of Economics and Centre for the Economics of Education. Anna Vignoles is a Senior Lecturer in the Economics of Education at the Institute of Education and a Research Fellow at the Centre for the Economics of Education.

1 Introduction

In 1998, up front tuition fees were introduced for degree courses in the UK. Although poorer students are exempt from such fees, or at least pay lower amounts, many commentators predicted that fees would cause a fall off in the number of students entering HE. Furthermore, student maintenance grants were reduced, and then abolished¹ the following year (1999). Critics argued that this de facto increase in the costs of attending university would reduce the number of *applicants* to HE² and that students from the poorest groups in society would be most likely to be put off, assuming that poorer students tend to be more debt and risk averse, and credit constrained (Callender (2003)).

On the other hand, many economists have argued that, quite apart from the efficiency and equity arguments in favour of the introduction of tuition fees for HE (see for example Barr and Crawford (1998); Dolton, Greenaway and Vignoles (1997); Greenaway and Haynes (2003)), the huge wage gains from a degree, combined with a relatively low tuition fee, would be unlikely to put students off going to university. Crudely put, given estimates of the wage gain from having a degree of just under 30 per cent, as compared to having just A levels (Dearden et al. (2002); Sianesi (2003)), it would perhaps be surprising if a £1000pa tuition fee were to dissuade large numbers from attaining graduate status³. Furthermore, the poorest students are exempt from fees anyway.

In any case, it has been argued that the problem of access to HE is in fact not rooted in the HE sector itself. It may be that the gap in educational achievement between poorer and richer students occurs prior to entry into university, i.e. poorer students may be much less likely to get good GCSEs and A-levels in the first place. If this is the case, it is conceivable that tuition fees might actually *reduce* inequality in the system. An

¹ Limited student grants were re-introduced in September 2002 for Welsh students. These are means tested and are for up to £1,500pa.

² Given that the supply of HE places is constrained, it is possible for student demand to fall whilst overall student numbers do not. In effect the excess demand for HE might have been reduced by the introduction of tuition fees.

³ See Goodman (2004) for an evaluation of the present value of a degree under the current financial arrangements.

argument can be made that if the access problem comes before entry into HE, diverting taxpayers' money away from HE towards earlier stages of our education system could *potentially* broaden access to higher education.

This paper provides up to date empirical evidence on the socio-economic gap in HE participation, for the period spanning the introduction of tuition fees. We assess whether the gap has widened and ask whether the socio-economic gap emerges on entry into university or much earlier in the education system. We do this in two ways. Firstly we consider the likelihood of going to university if you come from a poor neighbourhood and analyse changes in this likelihood over time. Secondly, we use more detailed individual level data to model the determinants of HE participation, focusing on changes in the relationship between family background and HE participation over time.

The paper relates to a burgeoning economic theoretical literature on educational inequality (e.g. Benabou 1996; Fernández and Rogerson 1996; Fernández and Rogerson 1998), as well as the sociological literature on this issue (e.g. Breen and Goldthorpe (1997)). The paper contributes, from an economic perspective, to the growing number of empirical studies that have investigated the factors that influence HE participation in the UK. Much of this empirical work has been from a sociological perspective and is summarised in Gayle et al (2003). Indeed Gayle et al. (2000; 2002 and 2003) have looked at this issue using different cohorts of the Youth Cohort Study (which is one of the data sets used in this paper) and their results are discussed at length in section 6 below, which presents our own results using different cohorts of this data set.

The paper is also specifically motivated by recent empirical evidence on educational inequality in the UK⁴, which suggests that between the late 1970s and the early 1990s, parental income became a more important determinant of whether an individual went on to higher education. Blanden and Machin (2003) have investigated in some detail the relationship between parental income and higher education participation. They conclude that the expansion of the education system in the 1970s, through to the

⁴ Blanden et al. (2002), Galindo-Rueda and Vignoles (forthcoming), Machin and Vignoles (forthcoming).

1990s, was associated with a widening of the gap in HE participation between rich and poor children. Glennerster (2001) also found evidence of a strengthening of the relationship between social class and the HE participation rate in the 1990s. His data ends however in 1998, which of course is when tuition fees were introduced⁵. We focus specifically on the period spanning the introduction of tuition fees in 1998, and on the relationship between social class and HE participation⁶, as well as parental income.

The next section gives more detail about the recent changes to higher education policy in the UK. Section 3 discusses the trends in UK higher education participation. In section 4, we discuss our methodology and data. Some neighbourhood evidence on changes in HE participation is considered in section 5. In section 6 we estimate our individual level models of the probability of studying for a degree. Finally, section 7 concludes.

2 Higher Education Policy in the UK

In recent decades, a number of countries have introduced tuition fees for higher education on efficiency and equity grounds (Woodhall (2002)). In the UK, tuition fees were implemented in 1998. Prior to that time, students did not pay for their higher education courses and there was a means tested grant. Tuition fees are currently payable up front and stand at £1,125 per annum. Students whose parents earn more than around £30,000pa are liable for the full amount. Some exemption is given for students whose parents earn between approximately £21,000 and £30,000pa, and students whose parents earn less than £21,000pa are exempt from fees. Student loans have also replaced the grant system. However, it is worth stressing that the real value of student grants was being eroded well before their abolition in the late 1990s.

⁵ Gayle et al. (2003) find no evidence of an increasing socio-economic gap in HE participation between Cohort 3 and Cohort 9 of the YCS but they acknowledge that quantitative comparisons over time using their results are not straightforward.

⁶ Erickson and Goldthorpe (1992), Erikson and Goldthorpe (1985), Saunders (1997) and Schoon et al. (2002) have examined issues relating to education and social mobility, to cite just a few.

In 2004, the UK parliament narrowly passed legislation to make further changes to the funding of higher education. Variable tuition fees have been proposed, i.e. fees that vary both by course and by institution. But perhaps the most important feature of the White Paper proposals, from the perspective of widening access, is that fees will be repaid after graduation via an income contingent loan system, and grants will be restored to low-income students. The empirical evidence presented here, which focuses on the period spanning the introduction of *up front* tuition fees, will not therefore necessarily apply to this new ICL scheme due to be introduced in 2005.

3 Trends in Higher Education Participation

The introduction of tuition fees does not appear to have had a major impact on aggregate HE participation⁷. Figures 1 and 2 show the simple trend of first-degree students (all years including new entrants and full and part-time students)⁸, by gender and over time, for England and Wales, and Scotland respectively⁹. A slight stagnation of the upward trend in student numbers is evident in all countries, following the introduction of tuition fees. However, the trend then resumes its upward path.

Access to HE in the UK has always been predominantly limited to those from higher socio-economic groups. Certainly if one looks at the very top and bottom of the socio-economic scale, the situation is dire. More than three quarters of students from professional backgrounds study for a degree, compared to just 14 per cent of those from unskilled backgrounds. Moreover, this inequality in the HE system has persisted over the last forty years.

⁷ From the Higher Education Statistics Agency.

⁸ We also excluded overseas students, those who did not report a domicile postcode, students with missing data on various fields, namely previous qualifications, institution type, qualification aim and degree subject. Full details of these samples are available from the authors. Figures 1 and 2 include students of all ages. For the main model, we restrict our sample to 18-24 year old students enrolled in HE.

⁹ Scotland is shown separately because both the educational system and the funding regime are different. In Scotland in September 2000, annual tuition fees were replaced by a £2,000 graduate contribution repayable after graduation.

An analysis of the DfES age participation index¹⁰ (Table 1) suggests a rise in participation by all socio-economic groups over the 1990s and a marginal widening of the gap in participation rates between richer and poorer students. The gap in the participation by the highest (A-C1) and lowest (C2-E) social groups is around 30 percentage points in 1996, prior to the introduction of fees, and 31 percentage points in 2001. These raw data do show some changes around the time of the introduction of tuition fees. There was a noticeable fall in participation in 1997 and 1998 (the former, possibly, in anticipation of the fees), before participation started rising again in 1999.

4 Methodology and Data

From a methodological perspective, determining the causal impact of tuition fees on the demand for HE is problematic. Up front tuition fees were introduced universally across England and Wales (and simultaneously student maintenance grants abolished). There was no “experiment” to determine the impact of fees, for example by introducing tuition fees in some areas but not others¹¹. Simply looking at student numbers before and after the introduction of tuition fees is likely to be informative but quite problematic, given that there has been a secular rise in the number of HE entrants over the last 30 years. We use two sources of data to address this issue in a largely descriptive way.

4.1 Higher education statistics agency data

The time series data we use comes from the Higher Education Statistics Agency, and covers the period 1994-2001. The data set includes limited information (gender, ethnicity, university, degree subject, home postcode, etc.) on all students in HE. We focus specifically on 18-24 year olds enrolled (either part time or full time) in HE. As

¹⁰ Based on the DfES Age Participation Index which measures the proportion of the under 21s in each social class participating in Higher Education for the first time (i.e. young entrants from each social class as a percentage of all young people in each social class).

¹¹ Tuition fees were not introduced in Scotland but there were other changes to the funding regime in that country making inter-country comparisons to evaluate the impact of fees on HE participation problematic. However, our analysis of HESA data suggested that the trends observed for England and Wales were broadly similar for Scotland too, and thus could not be explained by tuition fees specifically.

we are most interested in the impact of tuition fees on the participation of poorer students, and as the HESA data set does not contain information on the income or social class of the student's parents, we use the student's home postcode sector¹² as an indicator of their socio-economic status. This postcode sector is almost always the postcode of the students' parents and thus reflects the neighbourhood in which the person grew up.

Specifically, we have merged CACI Paycheck household income data into the HESA database to construct a postcode level dataset, on the basis of each student's home postcode. CACI data is derived from a commercially produced data set, based on over 4 million households¹³. This data set can provide us with an estimate of the income distribution of each postcode sector¹⁴. Furthermore, for each postcode sector we obtained from the 2001 Census the proportion of heads of household (age thirty five years or older) who classify themselves as being in socio-economic group I or II (professional or intermediate) and socio-economic group III (supervisory, clerical, junior managerial or administrative). This gives us separate indicators of the socio-economic profile of each neighbourhood.

Obviously the number of young people actually in higher education in each postcode will vary according to the total population of young people in each postcode. Obtaining annual estimates of the number of young people in each postcode proved impossible. We therefore have to rely on various population estimates from one point in time, namely the 2001 Census. This ignores any year on year population changes in each postcode. Two main population estimates were used. Firstly, we used the number of 18-24 year olds living in a particular postcode. However, neighbourhoods located near Colleges and Universities will tend to have a large 18-24 year old population, even if most of these students are not from that particular postcode originally. We therefore also used Census data on the number of 10-15 year olds living in each postcode as an alternative estimate of the number of young people in the postcode who are of the right age to enter HE.

¹² For ease of exposition, thereafter we refer to these postcode sectors as postcodes or neighbourhoods.

¹³ Further details, and an alternative use for these data, can be found in Gibbons (2001).

¹⁴ We grouped the information at postcode sector level as CACI provides data for the first 5 digits of the postcode only.

Another issue is that our indicators of the socio-economic composition of each postcode are not time varying. For example, we only have income data for each postcode for two years (1996 and 1999). We use the more recent 1999 data and have to assume some stability in the income distribution across different postcodes over the 8-year period. Comparisons of the 1996 and 1999 CACI data suggest that this assumption is reasonable. We will however, have introduced measurement error into an explanatory variable (the income level of the neighbourhood)¹⁵. The mean neighbourhood household income level is 21.89 (i.e. £21,890 pa in 1999 prices) with a standard deviation of 5.6 (£5,600). The mean proportion of heads of household that are in groups I/II is 20 per cent (s. d. 0.089) and in group III is 30 per cent (s. d. 0.054).

4.2 Youth cohort study data

The second data set we use for our micro analysis is the Youth Cohort Study (YCS), which is a series of longitudinal surveys conducted by the Department for Education and Skills. The surveys are of a particular academic year group or “cohort”, and are carried out contacting cohort members by post three times, at yearly intervals, when they were 16-17, 17-18 and 18-19. Respondents are first surveyed in the year after they are eligible to leave compulsory schooling. They are then followed up, generally over a two-year period¹⁶. The data collected includes information about the economic status of the young person, and in particular whether they have entered higher education by age 18/19, as well as their educational background, qualifications, family background and other socio-economic indicators. The survey is nationally representative (England and Wales) and the sample size of each cohort is around 20,000.

To date ten cohorts of young people have been surveyed. We concentrate on cohort 7, consisting of individuals who were aged 18 in 1996, and cohort 9, who were aged 18 in 2000. This spans the period during which tuition fees were introduced.

¹⁵ We have also dropped misreported or discontinued postcodes, constituting 3.05 per cent of the sample. Where we have missing data on postcode income levels, we use 1996 income data if available or, if that is missing too, we aggregate the data up to the 4-digit postcode level and impute data on the basis of this more aggregated grouping.

¹⁶ Some of the early cohorts have since been followed up to age 21 and beyond.

The YCS has been used extensively as a resource to analyse educational outcomes and subsequent transitions into the labour market (Croxford (2000); Dolton et al. (2001); Gayle et al. (2000; 2002; 2003), Howieson and Payne et al. (1996), Payne (1996); Rice (1999)). It is not without its faults however. It lacks measures of parental income, and thus we focus on the impact of parental social class¹⁷ as our main proxy indicator for family background (along with parental education). Attrition is the major problem in the YCS, and there has been extensive academic research on this issue (Lynn (1996)). Cohort 9 started out with an initial target sample size of 22,500. In the first survey at age 16/17, the response rate was 65 per cent. A similar response rate was also achieved in the 17-year-old and 18-year-old surveys. This means that the 18-year-old sample constitutes only 28 per cent of the initial sample (6,304 young people). Furthermore, the attrition rate is considerably higher in the Cohort 9 age 18 survey, than in the Cohort 7 age 18 survey, leading to a smaller sample size for Cohort 9. This is an issue we return to later when we present the descriptive statistics for the YCS samples and discuss our results. Here we simply note that the data are reweighed to allow for attrition and to bring them in line with population estimates, and our results are robust to re-weighting.

5 Neighbourhood Income Levels and Changes in HE Participation

We turn now to our postcode analyses. We want to explain why more young people from one postcode sector enter HE, as compared to other postcode sectors, and how this has changed over the period. In particular we want to assess more pupils from richer postcodes are enrolled in HE and whether the relationship between postcode income and HE participation has changed over time.

Table 2 presents various regression models that seek to explain differences in HE participation across postcodes and time. The dependent variable is the natural logarithm of the number of 18-24 year olds¹⁸ enrolled in HE. Here we only report the

¹⁷ The YCS provides data on the socio-economic group to which the individual belongs, based on his or her parents' occupation. We use the term social class for these groupings for ease of exposition.

¹⁸ As has been previously said, we only include students domiciled in England and Wales but we do

coefficients on postcode income measures, since our focus is on the nature of the relationship between neighbourhood income and the number of students going into HE¹⁹. However, the model also controls for other variables depending on the specification, as we now discuss.

All of the specifications include time dummies (base case is 1994/1995), allowing for national trends in HE participation over time. In Column 1 the only other control is the population estimate of the number of 18-24 year olds living in each postcode. Thus we allow for the fact that postcodes with more young people are likely to have more of them enrolled in HE. Column 2 includes the same variables but using a fixed effects formulation, thus focusing on changes within postcodes over time. This should remove differences in HE participation across postcodes that are down to characteristics of the postcode that we do not observe and are not explicitly included in our model.

Column 3 then includes interactions between the number of 18-24 year olds in the postcode and each year variable. Thus this specification tests whether ‘larger’ postcodes, i.e. those with more young people, experience a different trend in HE participation over time, as compared to smaller postcodes. In other words, it is an attempt to go some way to overcome the problem we mentioned earlier, namely that we do not have annual estimates of the population in each postcode.

Column 4 then adds terms indicating the qualification level in each postcode, i.e. the proportion of population in each postcode with a level 4 or 5 qualification (degree or above). This is to determine whether any positive relationship between neighbourhood income and HE participation is simply down to the fact that richer postcodes have higher proportions of more educated individuals who may be more likely to encourage their children to go on to HE.

Lastly, to check the robustness of the results, Column 5 uses an alternative measure of population, namely the number of 10-15 year olds in each postcode.

include all full time and part time students.

¹⁹ Full results available on request.

The basic OLS model in Column 1 suggests higher income postcodes have a higher number of young people enrolled in HE, as one would expect. What are of greater interest however, from a policy perspective, are the interactions between the income and year variables. Have richer postcodes experienced a more rapid increase in the number of young people going into HE, as compared to poorer neighbourhoods? These interaction terms are generally insignificant in this OLS model. In Column 2 however, once one allows for other differences across postcodes by using a fixed effects model, the interaction terms become always positive and significant. There is an increase in the coefficients by and large through to 1999, with some reduction thereafter. This suggests that for the early part of the period, richer neighbourhoods experienced a somewhat more rapid increase in the number of 18-24 year olds enrolled in HE over the period.

This same pattern is maintained in the other specifications, i.e. even when one allows for interactions between the number of young people in a postcode and the time dummies, and the qualification levels in each postcode. Using alternative measures of the population of young people (the number of 10-15 year olds in each postcode) does reduce the magnitude of the coefficients somewhat but still suggests that postcode income level was more positively associated with HE participation in 1996 through to 2000, as compared to 1994. We also undertook a similar analysis using a different measure of the socio-economic profile of the neighbourhood, namely the proportion of heads of household who classify themselves as being in socio-economic group I or II (professional or intermediate) and the proportion in socio-economic group III. These results gave qualitatively similar results²⁰.

This can be shown more clearly in Figure 3. For this figure, we calculated the ratio of the number of students enrolled in HE to the number of 18-24 year olds in that postcode. We then show the trend in this ratio for postcodes from the top, middle and bottom quintiles of the postcode household income distribution. The gap in the ratio for rich, middle and poor postcodes is normalised to zero for 1994. Figure 3 then clearly shows the widening gap in HE participation between postcodes from the top household income quintile and those from the middle or bottom quintiles. This

²⁰ Results available on request.

widening of the gap pre-dates the introduction of fees and appears to be part of a longer-term trend dating back at least to the early 1990s. Figures 4 and 5 use the same methodology to analyse separately the number of students enrolled in ‘new’ and ‘old’ universities²¹. A similar pattern is observed for both types of institution, however the gap between rich and poor postcodes is greater for the ‘old’ universities. We know that students from poorer neighbourhoods are not only less likely to go into HE in the first place but also make different choices of institution. Our research suggests that HE participation in richer neighbourhoods increased more rapidly during the 1990s as compared to HE participation in poorer neighbourhoods, and that this trend was particularly acute for those choosing to attend ‘old’ universities. Choice of institution by different types of student is certainly an issue that merits further research.

We undertook a number of further robustness checks. Firstly, we used different measures of postcode income. For example, we used the ranking of the postcode in the income distribution, which should abstract from any changes in the shape of the income distribution over this time period. We also used income quintiles to indicate the income level of each postcode, as in Figure 3. Both these methods yielded qualitatively similar results.

The timing of these trends suggests that causes other than tuition fees may well be responsible. The continuing decline in the real value of student grants might be one possible culprit. However, in this analysis we cannot talk about causality. For example, we do not control for students’ prior attainment (at area level) and thus we cannot be sure whether we are observing increasing socio-economic inequality on entry into HE or the results of increasing inequality emerging far earlier in the education system. Therefore, we now turn to our micro-analysis of the YCS data in an attempt to identify more precisely the determinants of HE participation over the period in question.

²¹ Using the *Guardian* newspaper classification of universities into old and new, available on request.

6 Micro-analysis of the Determinants of HE Participation

Descriptive statistics for the YCS sample are shown in Table 3. The first columns describe the samples participating in higher education and not participating in HE for the cohort aged 18 in 1996. Of those aged 18 in 1996, 32.8 per cent were in higher education. The second set of columns provides the same information for the cohort aged 18 in 2000, of which 40 per cent were participating in higher education.

Even over this relatively short period of time, there have been some changes in the characteristics of those participating in higher education. Of those participating in HE in 1996, around one third were from a professional, managerial or technical background. By 2000, this had risen to 42 per cent. Over the period there was a decline in the proportion of HE students from skilled manual, semi-skilled and unskilled backgrounds. Some of the changes observed are very substantial. 8 per cent of those participating in HE in 1996 had a father with a degree: by 2000 this had risen to 14 per cent, and a similar trend is observed for the proportion of students whose mother had a degree. Whilst 55 per cent of those participating in HE in 1996 had 3 or more A levels, this proportion had risen to 72 per cent for those in HE in 2000. Of course this latter trend may reflect rising A level achievement across the board, as much as a change in the composition of the HE student body. We suspect however, that the higher attrition rate for Cohort 9 may also explain some of the changes we observe. We discuss how we deal with this problem below, when we present the results of our modelling.

Table 4 shows the marginal effects from a probit model, where the dependent variable took a value of one if the person was in higher education at age 18 and zero if they were not²². Individuals not participating in HE could be in various states, either in or out of the labour market, or studying for lower level qualifications. The model is estimated separately for the two cohorts. Specification 1 shows the impact on HE participation of gender, socio-economic background, ethnicity, parental education and school type. The purpose of this model is to measure the maximum possible impact

²² Thus, we are measuring HE participation not degree attainment. If drop out and degree failure vary by social class, and if these variables have been affected by the introduction of tuition fees, we may be underestimating the impact of social class on HE achievement.

from an individual's socio-economic background on the likelihood of HE participation and we therefore do not include the individual's level of prior academic achievement. In this model, family background may act directly on the decision to enter HE at 18 or indirectly via lower academic achievement prior to HE. It is nonetheless an important policy question, as to whether family background impacts on pupils' educational outcomes earlier in the education system or only on entry into HE²³. We test this argument in specification 2 by including measures of the student's prior academic achievement, namely the number of A levels held.

Our results from Table 4 show quite clearly that a students' socio-economic background has a large impact on the probability of participating in HE when one does not control for prior educational achievement. Thus in column 1, the results suggest that students from a professional/ managerial or technical background are almost 3 percentage points more likely to be participating in HE in 1996. The impact from this variable disappears once indicators of the students' achievement at A level are included. In other words, in 1996, for a given level of achievement at age 18, socio-economic background does not appear to have an additional impact on the decision to go to university.

By 2000 however, the situation appears to have changed somewhat. Not only is the impact of coming from a professional/ managerial or technical background larger (12 percentage points) but it remains sizeable (6 percentage points) and significant even when A level measures are included in the model. A similar pattern is observed for students with non-manual backgrounds.

At the other end of the distribution, in 1996, students with unskilled parents were not significantly less likely to attend HE compared to students from a skilled background. By 2000, students from an unskilled background were ten percentage points less likely to attend HE, not controlling for prior attainment. Once achievement at A level was included in the model however, the negative impact of coming from an unskilled background became insignificant.

²³ There is a large literature relating to when family background has the greatest impact on pupil attainment at school and the optimal timing of any policy interventions to improve the educational attainment of poor children (Cameron and Heckman (2001); Carneiro et al. (forthcoming); Currie and Thomas (1999); Haveman and Wolfe (1995)).

The impacts of other family background characteristics are also of interest, although we cannot comment on them all for reasons of space²⁴. Noteworthy changes include a small decline in the impact of parental education on HE participation, and a reduction in the impact from attending a selective (grammar) school. Of course one could argue that school type is endogenous, reflecting pupil prior attainment and choices made by different pupils and families. In specifications without the school type variables, the impact of socio-economic background increases, as one would expect. However, the general pattern of a large increase over the period in the impact from coming from a higher socio-economic background on HE participation remains with or without these additional controls. Moving to the specifications that control for attainment at A level, it appears that the impact of qualifications attained prior to entry to HE became only marginally more important in 2000 than in 1996.

We conclude that over the period there was an increase in the impact of the social class variables and that this increased impact remained even after controlling for A level attainment and school type. There is however a potential problem with these results. As we have discussed, the sample from Cohort 7 is substantially larger by age 18 than is the case for Cohort 9, reflecting the different attrition rate across the two cohorts. Thus the rather significant changes in the coefficients (in particular the rise in the magnitude of the coefficients on the socio-economic background variables and reduction in the importance of parental education) may partly reflect attrition bias caused by students from poorer backgrounds attriting to a greater extent in the Cohort 9 survey. There is some evidence to support this in the descriptive statistics. One possible solution to this problem is to restrict the sample to a more homogenous group of higher attaining pupils, where the attrition problem is less of an issue. We therefore re-estimate our models, restricting the sample to only those who achieved 5 good GCSEs, as discussed below.

6.1 GCSE comparator group

²⁴ For example, our results are robust to the inclusion of regional fixed effects. Our sample size is not sufficient to estimate the model by region but we are aware of very recent evidence of regional differences in the determinants of education participation (Rice (2004)).

Of course a legitimate question is whether the comparator group used for the analysis in Table 4 is appropriate anyway. It could be argued that we should restrict our comparator group to those who are potentially able and qualified to go on to HE²⁵. Restricting our sample to those with five or more good General Certificates of Secondary Education (GCSE) would appear to be appropriate²⁶. Increasingly individuals enter HE without A levels and thus limiting the sample to only those with A levels is perhaps too restrictive. However, our results remain qualitatively similar even when the comparator group is all individuals with one or more A levels. Table 5 presents a summary table of the results from this more restricted sample²⁷, however before discussing these we need to address the problems of potential ability bias.

One of the weaknesses of the YCS data, as compared to other British cohort data sets, is that we have no independent measures of ability. Thus it is possible to argue that our socio-economic background variables are subject to ability bias. To test this we re-estimated our specifications including proxy indicators of ability, namely GCSE mathematics and English grade²⁸, as shown in specification 3 in Table 5²⁹. We recognise that GCSE grades are potentially endogenous variables, as for that matter are our measures of attainment at A level. Students may well not apply as much effort to their GCSE studies if they do not expect to stay on at school and take A levels. Likewise students may not take many A levels if they do not expect to go on to HE³⁰. However, from a policy perspective it is useful to have an indication when the socio-economic inequality in the education system emerges. By controlling for GCSE grade score, we can ask the question whether, for a given level of achievement at GCSE and

²⁵ There is potential sample selection bias here. The group of individuals with good GCSEs (and indeed those studying for A-levels) are not representative of the cohort as a whole. To estimate a sample selection model to allow for this, one would have to find a variable that can predict the likelihood of getting a good set of GCSEs, but does not impact subsequently on the likelihood of entering HE. We were unable to find such a variable (see Gayle et al. (2003) for a sample selection model using YCS).

²⁶ Appendix A shows the results of a probit model estimating the determinants of getting 5 or more good GCSEs. The impact of parental socio-economic status became more important between 1996 and 2000. This suggests that success in the education system at GCSE was also becoming more closely linked to family background (see Blanden et al. (2003) and Galindo-Rueda and Vignoles (2004) for a further discussion of this trend). Rice (2004) has also conducted a comprehensive analysis of the determinants of staying on at age 16 using YCS data for the whole of the 1990s. See also Gayle et al. (2002).

²⁷ Full results available on request.

²⁸ Coded as Grade A – 5 points, Grade B – 4 points, Grade C – 3 points, Grade D – 2 points, Grade E – 1 point. Otherwise zero.

²⁹ These variables were also added to the models on the full sample- results available on request.

³⁰ Since most people who take at least one A level go on to university, this measure may be somewhat less problematic.

attainment at A level, students from lower socio-economic groups are less likely to participate in HE³¹.

Table 5 confirms the pattern of results in Table 4, namely that we cannot find significant social class effects in 1996 but by 2000 some coefficients on the socio-economic status variables become positive and significant. However, when we add in our (endogenous) measures of GCSE grade performance, the impact of the social class variables becomes smaller and insignificant, even in 2000. As in Table 4, the effect of attainment at A level became somewhat more important by 2000, even when the comparator group is more restricted and even in the specifications that control for GCSE grades. The GCSE grade variables themselves are highly positively significant, as one would expect. However, there is a small diminution of the effect from both mathematics and English GCSE grades over the period.

It is worth noting that our results are largely consistent with other recent work on this data set, particularly by Gayle et al. (2003). Gayle et al. (2003) investigated the range of different factors that influence higher education participation using Cohort 9 of the YCS, i.e. the cohort that turned eighteen in 1999. Their main conclusions were that prior qualifications, parental education, region and housing tenure are the key factors influencing HE participation. They report that comparisons over time, i.e. with their earlier work on Cohort 3 (age 18 in 1989), are quite problematic due to changes in survey questions etc. However, they found some evidence that between Cohort 3 and Cohort 9, the gender gap in HE participation had been closed. On the other hand, they suggested that the socio-economic gap in HE participation had not been closed during this period. They found, as we did for both our YCS cohorts, that prior attainment is the crucially important determinant of HE participation and that socio-economic background has little affect once one adequately controls for attainment at GCSE and A level. As we do, they conclude that resources might better be directed towards earlier phases of the education system in order to close or reduce the socio-economic gap.

³¹ We are currently devising an instrument for GCSE attainment at age 16, based on the average GCSE grade of the pupils in the individual's school.

6.2 Changes over time

The previous section suggested that there has been increasing income-driven inequality in HE participation over a longer period of time, i.e. from 1994/5 onwards, and not simply after the introduction of tuition fees. This is supported by other evidence of a long-term increase in income-driven inequality in HE in the UK (Blanden and Machin (2004); Machin and Vignoles (forthcoming)). Blanden and Machin (2004) find that, even after controlling for attainment at A level, there was an increasing effect from parental income on HE participation from the 1970s to the 1990s. Our YCS results by contrast do not indicate an independent effect from socio-economic background on HE participation, conditional on earlier educational attainment, even by 1996. However, since our YCS models assess the impact of socio-economic background, whilst Blanden and Machin use parental income, comparisons between the two pieces of research are problematic. It is quite conceivable that the relatively small impact from parental income found by Blanden and Machin (2004) is not being picked up by our broader measures of family background.

7 Conclusions

Children from all socio-economic backgrounds are considerably more likely to go to university in 2001, as compared to 1994. In fact the growth in HE participation amongst poorer students has been remarkably high, mainly because they were starting from such a low base. Nonetheless our results suggest that poorer neighbourhoods (postcodes) saw a less rapid growth in the number of young people enrolled in HE as compared to richer neighbourhoods, particularly in the early and mid 1990s. The strength of the relationship between neighbourhood income levels and HE enrolment grew most rapidly in the early part of the period, rather than after the introduction of tuition fees. This would seem to imply that any income-driven inequality in HE is part of a longer-term trend, perhaps related to the gradual reduction in student support in HE and the big expansion of the university sector that occurred in the early 1990s³².

³² A number of policy changes occurred at this time, most of which caused student numbers in HE to

Our detailed individual level analysis suggests that in 1996, i.e. before tuition fees, there was certainly substantial social class educational inequality in HE but that it occurred largely as a result of inequalities earlier in the education system. By 2000 however, one can observe social class effects on HE participation, even after conditioning for the number of GCSEs and A levels that an individual has. This seems to suggest a widening of the social class gap in higher education itself in the period after the introduction of tuition fees. However, in models that include finer measures of educational achievement, the social class effects become smaller and insignificant. We conclude that much of the impact from social class on university attendance actually occurs well before entry into HE. Of course just because we observe inequality in attainment at earlier ages, does not mean it is not related to problems in HE. Students may look forward and anticipate barriers to participation in HE and make less effort in school as a result. Indeed there are many such potential barriers, not least of which is the expected cost of HE (see for example Connor et al. (2001)). This is an area that requires still further research.

grow. One significant policy change was the merging of 'old' universities and polytechnics into one new category just entitled 'universities'. A disproportionate amount of the growth in HE in the early 1990s occurred in former polytechnics. The same patterns can be observed however whether one considers HE participation in old or new universities (evidence on this is available on request).

References

Barr, N. and Crawford, I. (1998) 'The Dearing Report and the Government Response: A Critique', *Political Quarterly*, Vol. 69, No. 1, January - March, pp. 72-84.

Barr, N. (2002) 'Funding Higher Education: Policies for Access and Quality', House of Commons Education and Skills Committee, Post-16 student support, Session 2001-02, 24 April 2002.

Blanden, J and S. Machin (2003) 'Educational Inequality and the Expansion of UK Higher Education', Centre for Economic Performance mimeo.

Blanden, J., Gregg, P. and Machin, S. (2003) 'Changes in Educational Inequality', Centre for the Economics of Education Discussion Paper, forthcoming.

Blanden, J., Goodman, A., Gregg, P. and Machin, S. (2002) 'Changes in Intergenerational Mobility in Britain', Centre for the Economics of Education Discussion Paper No. 26, London School of Economics, forthcoming in M. Corak (ed.) *Generational Income Mobility in North America and Europe*, Cambridge University Press.

Breen, R. and Goldthorpe, J. (1997) 'Explaining Educational Differentials – Towards A Formal Rational Action Theory', *Rationality and Society*, Vol. 9, pp. 275-305.

Breen, R. and Goldthorpe, J. (1999) 'Class Inequality and Meritocracy: A Critique of Saunders and An Alternative Analysis', *British Journal of Sociology*, 50 (1), pp. 1-27.

Callender, C. (2003) 'Student Financial Support in Higher Education: Access and Exclusion', in M. Tight (ed.) *Access and Exclusion: International Perspectives on Higher Education Research*, Elsevier Science, London.

Cameron, S. and Heckman, J. (2001) 'The Dynamics of Educational Attainment of Black, Hispanic and White Males', *Journal of Political Economy*, 109, pp. 455-499.

Carneiro, J., Hansen, K. and Heckman, J. (2003 - forthcoming) 'Estimating Distribution of Counterfactuals with an Application to the Returns to Schooling and Measurement of the Effect of Uncertainty on Schooling Choice', *International Economic Review*.

Chevalier, A. and Lanot, G. (forthcoming) 'Financial Transfer and Educational Achievement', *Education Economics*.

Connor, H., Dewson, S., Tyers, C., Eccles, J. and Aston, J. (2001) *Social Class and Higher Education: Issues Affecting Decisions on Participation by Lower Social Class Groups*, Department for Education and Skills Research Report RR267.

Currie, J. and Thomas, D. (1999), 'Early Test Scores, Socioeconomic Status and Future Outcomes', NBER Working Paper No. 6943. National Bureau of Economic Research: Cambridge, Ma.

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

Dearden, L., McIntosh, S., Myck, M. and Vignoles, A. (2002) 'The Returns to Academic, Vocational and Basic Skills in Britain', *Bulletin of Economic Research*, vol. 54, no. 3, pp. 249-274.

Dolton, P., Greenaway, D. and Vignoles, A. (1997) 'Whither Higher Education? An Economic Perspective for the Dearing Committee of Inquiry', *Economic Journal*, vol. 107, No. 442, pp. 710-726.

Dolton, P., Makepeace, G. and Gannon, B. (2001) 'The earnings and employment effects of young people's vocational training in Britain', *Manchester School*, Vol. 69, No. 4, pp. 387-417.

Erickson, R. and Goldthorpe, J. (1992) 'The Constant Flux: A study of Class Mobility in Industrial Societies', Oxford University Press.

Galindo-Rueda, F. and Vignoles, A. (forthcoming) 'The Declining Relative Importance of Ability in Predicting Educational Attainment', *Journal of Human Resources* and forthcoming CEE Discussion Paper.

Gayle, V., Berridge, D. and Davies, R. (2003) *Econometric Analyses of the demand for Higher Education*, Department for Education and Skills Research Report RR471.

Gayle, V., Berridge, D. and Davies, R. (2002) 'Young People's Entry to Higher Education: Quantifying Influential Factors', *Oxford Review of Education*, Vol. 28, pp. 5-20.

Gayle, V., Berridge, D. and Davies, R. (2000) 'Young People's Routes to Higher Education: Exploring Social Processes with Longitudinal Data', *Higher Education Review*, Vol. 33, pp. 47-64.

Gibbons, S. (2001) 'Paying for Good Neighbours? Neighbourhood Deprivation and the Community Benefits of Education', Centre for the Economics of Education, Discussion Paper No. 17, London School of Economics.

Glennerster, H. (2001) 'United Kingdom Education 1997-2001', Centre for the Analysis of Social Exclusion (CASE), Paper 50.

Goodman, A. (2004) presentation to Department of Trade and Industry (DTI).

Greenaway, D. and Haynes, M. (2003) 'Funding Higher Education in the UK: The Role of Fees and Loans', *Economic Journal*, 113, F150-167.

Haveman, R. and Wolfe, B. (1995) 'The Determinants of Children's Attainments: A Review of Methods and Findings', *Journal of Economic Literature*, 33 (4), pp. 1829-1878.

Howieson, C. & Croxford, L. (1996) 'Using the YCS to analyse the outcomes of careers education & guidance'. The Stationary Office, DFEE Research Series, London.

Lynn, P. (1996) *England and Wales Youth Cohort Study: The Effect of Time Between Contacts, Questionnaire Length, Personalisation and Other Factors on Response to the YCS (Research Studies)*, London: The Stationery Office Books.

Payne, J. (2001) 'Patterns of Participation in Full-Time Education After 16: An analysis of the England and Wales Youth Cohort Studies', Nottingham, UK: Department for Education and Skills.

Payne, J., Cheng, Y. and Witherspoon, S. (1996) *Education and Training for 16-18 Year Olds in England and Wales – Individual paths and national trends*, Policy Studies Institute, London.

Machin, S. and Vignoles, A. (forthcoming) 'Education Inequality', *Fiscal Studies*, submitted March 2003.

Rice, P. (2004) 'Education and Training Post-16 - Differences across the British Regions'. CEE conference, March.

Rice, P (1999) 'The impact of local labour markets on investment in further education: evidence from the England and Wales youth cohort studies', *Journal of Population Economics*, Vol.12, No.2, pp.287- 31.

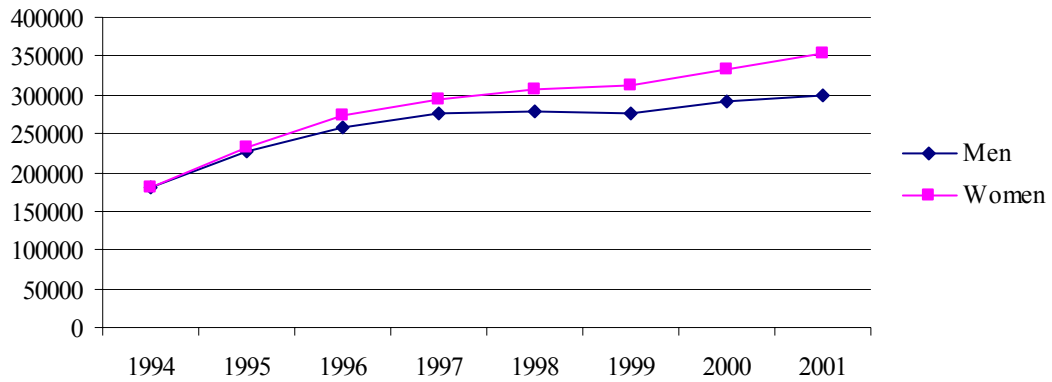
Saunders, P. (1997), 'Social Mobility in Britain: An Empirical Evaluation of Two Competing Explanations', *Sociology*, May, 31.

Sianesi, B. (2003) 'Returns to Education: A non-technical summary of CEE work and Policy Discussion', draft report for Department for Education and Skills.

Schoon, I., Bynner, J., Joshi, H., Parsons, S., Wiggins, RD. and Sacker, A. (2002) 'The influence of context, timing, and duration of risk experiences for the passage from childhood to mid-adulthood', *Child development*, Sep-Oct 2002, Vol. 73, No. 5, pp.1486-1504.

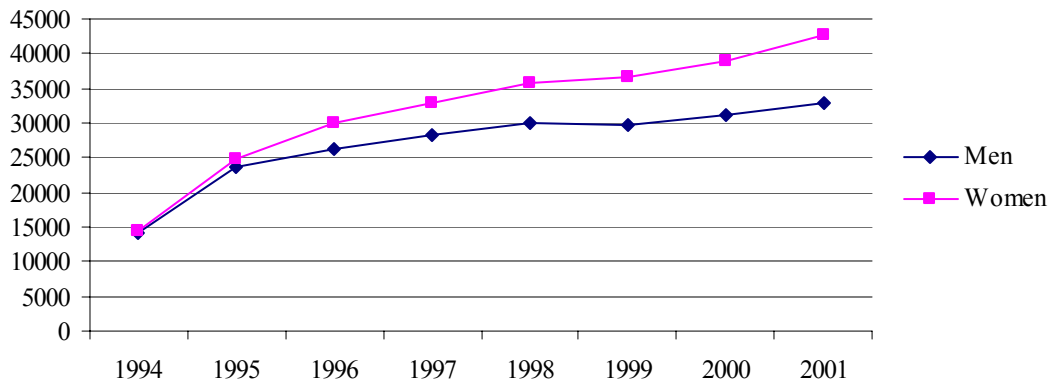
Woodhall, M (Ed) (2002) 'Paying for Learning: The Debate on Student Fees, Grants and Loans in International Perspective', *The Welsh Journal of Education*, Vol. 11, No. 1.

Figure 1: Total number of students studying for a first degree, by gender (England and Wales)



Source: Higher Education Statistics Agency.

Figure 2: Total number of students studying for a first degree, by gender (Scotland)



Source: Higher Education Statistics Agency.

Figure 3: Changes in postcode-level HE participation over time, by neighbourhood income level

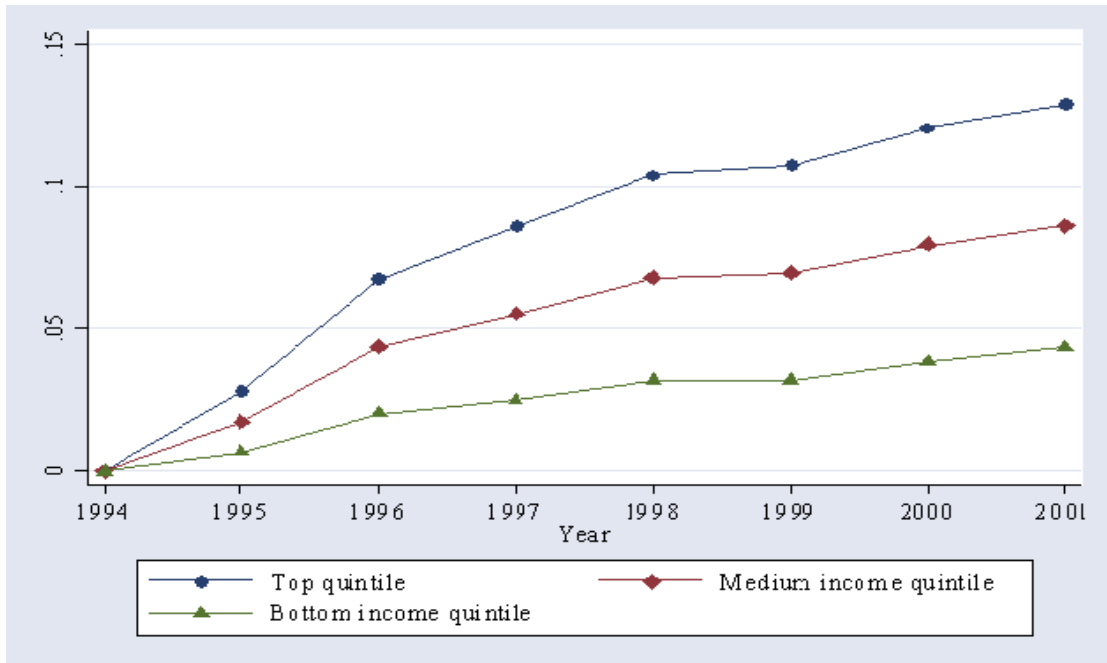


Figure 4: Changes in postcode-level HE participation over time (“new universities”), by neighbourhood income level

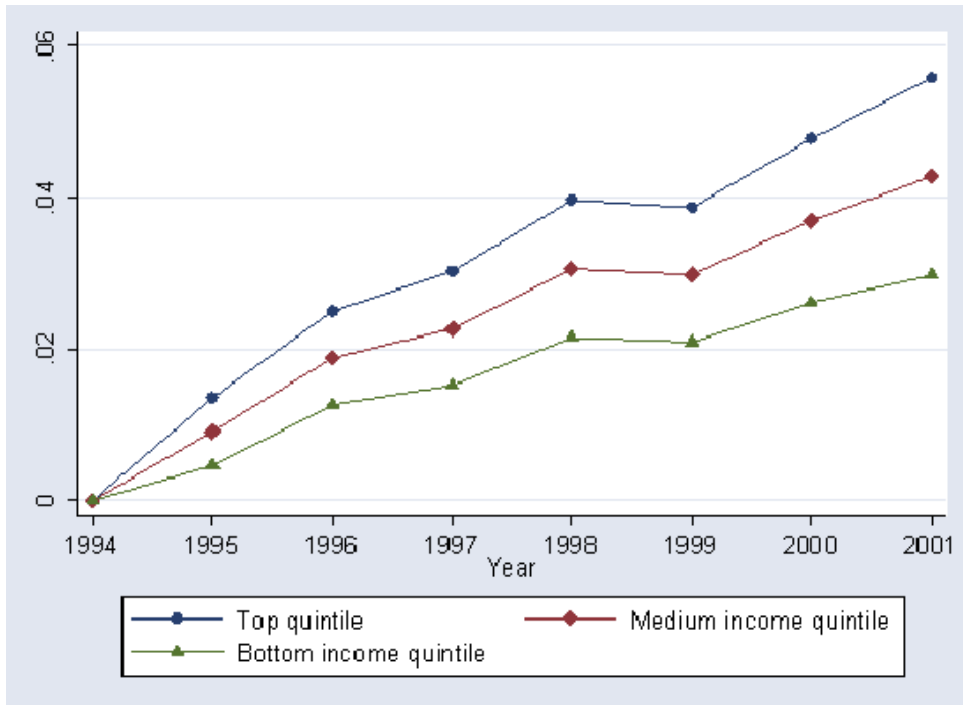


Figure 5: Changes in postcode-level HE participation over time (“old universities”), by neighbourhood income level

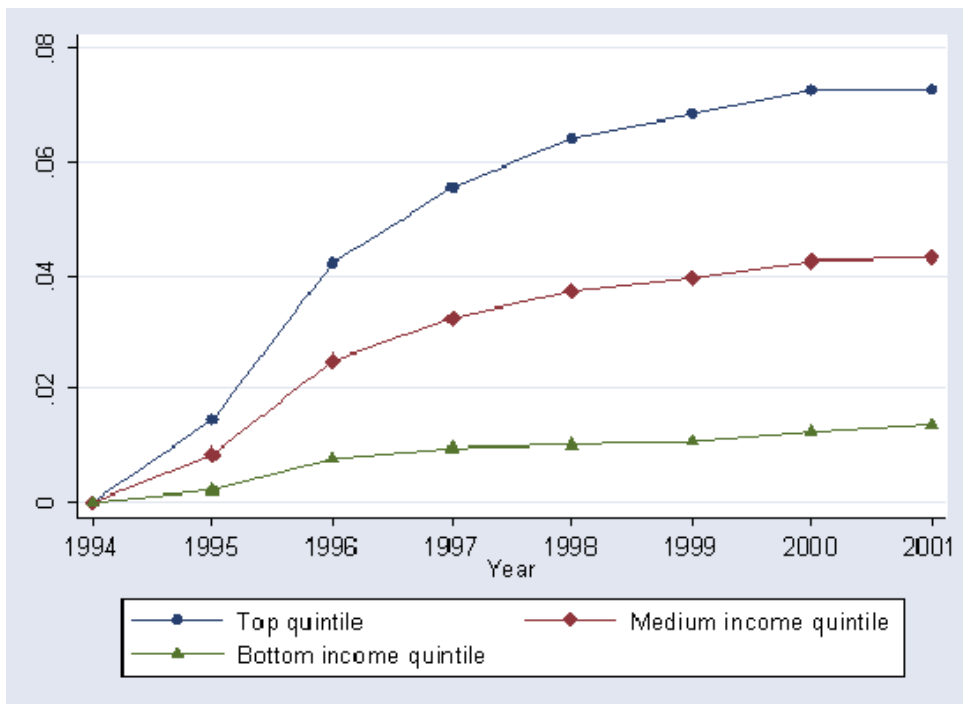


Table 1: Age Participation Index (API) (%) by social class, 1991/2 – 2001

Class	Year of Entry									
	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
Professional (A)	71	73	78	80	82	79	72	73	76	79
Intermediate (B)	39	42	45	46	47	48	45	45	48	50
Skilled non-man (C1)	27	29	31	31	32	31	29	30	33	33
Skilled manual (C2)	15	17	18	18	18	19	18	18	19	21
Partly skilled (D)	14	16	17	17	17	18	17	17	19	18
Unskilled (E)	9	11	11	12	13	14	13	13	14	15
A-C1	40	43	46	47	48	48	45	45	48	50
C2-E	14	16	17	17	18	18	17	17	18	19

Source: Department for Education and Skills Age Participation Index which measures the proportion of the under 21s in each social class participating in Higher Education for the first time (i.e. young entrants from each social class as a percentage of all young people in each social class).

Table 2: Postcode level analysis of the number of students attending HE

	Dependent Variable: Ln (Number of HE students) in each postcode				
	OLS	Fixed Effects	Fixed Effects	Fixed Effects	Fixed Effects
Ln (Income)	1.333 (36.95)***	-	-	-	-
Interaction Ln (Income)*Year 1995	0.096 (1.87)*	0.052 (2.67)***	0.061 (3.10)***	0.076 (2.71)***	0.033 (1.20)
Interaction Ln (Income)*Year 1996	0.075 (1.48)	0.096 (4.95)***	0.108 (5.60)***	0.165 (5.99)***	0.095 (3.49)***
Interaction Ln (Income)*Year 1997	0.047 (0.93)	0.099 (5.13)***	0.112 (5.78)***	0.139 (5.04)***	0.074 (2.74)***
Interaction Ln (Income)*Year 1998	0.029 (0.57)	0.082 (4.27)***	0.089 (4.60)***	0.081 (2.95)***	0.051 (1.89)*
Interaction Ln (Income)*Year 1999	0.071 (1.41)	0.142 (7.37)***	0.147 (7.62)***	0.136 (4.93)***	0.113 (4.18)***
Interaction Ln (Income)*Year 2000	0.027 (0.53)	0.103 (5.37)***	0.109 (5.65)***	0.082 (2.97)***	0.058 (2.14)**
Interaction Ln (Income)*Year 2001	0.017 (0.34)	0.080 (4.18)***	0.087 (4.53)***	0.058 (2.10)**	0.029 (1.06)
Controls:					
Year dummies	}	}	}	}	}
Population aged 18-24 from Census (in thousands)	}	}	}	}	}
Interactions Population aged 18-24*years			}	}	}
Ratio Population aged older 24 Qualification level 4/5				}	}
Interactions Ratio (Population qualification level 4/5/Population aged older 24)*Years				}	}
Population aged 10-15 from Census (in thousands)					}
Interactions Population aged 10-15*years					}
Constant	-1.046 (9.46)***	3.465 (997.21)***	3.465 (997.89)***	3.465 (998.12)***	3.465 (998.10)***
Observations	57743	57743	57743	57743	57743
R-squared	0.34	0.40	0.40	0.40	0.40
Number of postcodes	-	7382	7382	7382	7382

Data: Higher Education Statistics Agency. Sample restricted to full and part-time students age 18-24 year olds domiciled in England or Wales enrolled in a first-degree course. See main paper for other sample restrictions.

Table 3: Descriptive Statistics for YCS Sample

	1996		2000	
	Participating in HE	Not participating in HE	Participating in HE	Not participating in HE
Male	0.41	0.43	0.39	0.42
Parents' Socio-Economic Status:				
Professional, managerial & tech. oc.	0.29	0.20	0.38	0.22
Other non-manual occupations	0.21	0.18	0.27	0.21
Skilled occupations – manual*	0.31	0.34	0.23	0.33
Semi-skilled occupations - manual	0.08	0.12	0.07	0.11
Unskilled occupations	0.02	0.03	0.01	0.04
Other	0.08	0.13	0.04	0.08
Ethnicity:				
White*	0.92	0.92	0.90	0.91
Black	0.01	0.01	0.01	0.02
Asian	0.07	0.06	0.08	0.06
Other	0.01	0.01	0.01	0.02
Parental education:				
Father degree	0.07	0.03	0.12	0.09
Father at least one A level	0.32	0.15	0.32	0.15
Father below one A level*	0.61	0.81	0.56	0.76
Mother degree	0.04	0.02	0.16	0.11
Mother at least one A level	0.31	0.15	0.23	0.12
Mother below one A level*	0.65	0.83	0.61	0.78
Type of school attended:				
Comprehensive age 16	0.24	0.35	0.24	0.33
Comprehensive age 18*	0.49	0.53	0.51	0.54
Grammar	0.08	0.03	0.09	0.04
Secondary modern	0.02	0.04	0.01	0.04
Independent	0.17	0.05	0.15	0.06
Highest school qualification:				
One or two A levels	0.12	0.11	0.11	0.13
Three or more A levels	0.55	0.11	0.72	0.19
Five or more A-C GCSEs	0.96	0.47	0.98	0.57
Observations	2343	5642	2186	4082
	7985		6268	

Data: Youth Cohort Survey.

* Base case in subsequent regressions.

Table 4: The Determinants of HE Participation (Marginal Effects)

	1996		2000	
	Specification 1	Specification 2	Specification 1	Specification 2
Male	-0.008 (0.78)	0.000 (0.02)	-0.036 (2.91)***	-0.009 (0.70)
Parents' Socio-Economic Status:				
Professional, managerial & tech. occ.	0.027 (1.89)*	0.009 (0.59)	0.122 (6.72)***	0.060 (3.19)***
Other non-manual occ.	0.022 (1.50)	0.006 (0.43)	0.095 (5.23)***	0.046 (2.44)**
Semi-skilled occ.-manual	-0.032 (1.75)*	-0.029 (1.51)	-0.023 (0.96)	-0.024 (0.94)
Unskilled occ.	-0.008 (0.25)	0.009 (0.26)	-0.101 (2.60)***	-0.059 (1.33)
Other	-0.054 (3.00)***	-0.048 (2.58)***	-0.034 (1.15)	-0.039 (1.25)
Ethnicity:				
Black	-0.044 (0.77)	-0.021 (0.34)	-0.062 (1.17)	0.011 (0.19)
Asian	0.089 (3.86)***	0.109 (4.59)***	0.130 (5.00)***	0.164 (5.87)***
Other ethnicity	-0.057 (0.94)	-0.078 (1.30)	-0.065 (1.23)	-0.028 (0.49)
Ethnicity missing flag	-0.106 (1.78)*	-0.075 (1.16)	0.007 (0.10)	0.028 (0.36)
Parental education:				
Father degree	0.107 (3.76)***	0.047 (1.66)*	0.010 (0.46)	0.022 (0.95)
Father at least one A level	0.120 (7.94)***	0.069 (4.52)***	0.089 (4.52)***	0.063 (3.08)***
Father education missing flag	-0.012 (0.73)	0.002 (0.14)	-0.071 (3.90)***	-0.020 (1.03)
Mother degree	0.110 (3.05)***	0.060 (1.68)*	0.062 (3.11)***	0.038 (1.84)*
Mother at least one A level	0.096 (6.49)***	0.063 (4.16)***	0.073 (3.66)***	0.010 (0.48)
Mother education missing flag	-0.055 (3.43)***	-0.039 (2.38)**	-0.062 (3.31)***	-0.044 (2.21)**
Type of school attended:				
Comprehensive age 16	-0.040 (3.40)***	-0.033 (2.69)***	-0.038 (2.68)***	-0.026 (1.71)*
Grammar	0.251 (9.11)***	0.151 (5.45)***	0.131 (4.74)***	0.005 (0.19)
Secondary modern	-0.117 (4.09)***	-0.075 (2.45)**	-0.166 (4.41)***	-0.069 (1.60)
Independent	0.195 (9.72)***	0.171 (8.44)***	0.155 (6.71)***	0.113 (4.79)***
Highest school qualification:				
One or two A levels		0.190 (10.61)***		0.227 (10.63)***
Three or more A levels		0.488 (35.89)***		0.514 (35.18)***
-2 (log_R – log_U)	760.5***	2116.8***	653.4***	1998.4***
Observations		7985		6268

Data: Youth Cohort Survey Full Sample.

Notes: Dependent variable – value of one if in HE, zero otherwise, probit estimation. Base case: skilled background, white, father/ mother's education less than A level, attended a comprehensive and with no A levels.

Absolute value of t statistics in parentheses * significant at 10 per cent; ** significant at 5 per cent; *** significant at 1 per cent.

Table 5: The Determinants of HE Participation (Marginal effects) – Sample with 5+ Good GCSEs

	1996			2000		
	Specif. 1	Specif. 2	Specif. 3	Specif. 1	Specif. 2	Specif. 3
Parents' Socio-Economic Status:						
Professional, managerial & t.. occ.	0.013 (0.69)	0.004 (0.20)	-0.008 (0.39)	0.088 (4.06)***	0.052 (2.28)**	0.032 (1.36)
Other non-manual occ.	0.016 (0.77)	0.004 (0.18)	-0.008 (0.38)	0.074 (3.39)***	0.042 (1.81)*	0.024 (1.04)
Semi-skilled occ.-manual	-0.003 (0.11)	-0.009 (0.32)	-0.004 (0.14)	0.011 (0.35)	-0.009 (0.28)	-0.011 (0.32)
Unskilled occ.	0.044 (0.84)	0.054 (1.00)	0.056 (1.00)	-0.010 (0.17)	0.003 (0.04)	0.009 (0.14)
Other	-0.035 (1.28)	-0.034 (1.19)	-0.029 (1.01)	0.045 (1.08)	0.014 (0.31)	0.011 (0.25)
GCSE grades:						
GCSE maths grade			0.115 (13.10)***			0.074 (7.97)***
GCSE math grade missing flag			-0.138 (0.73)			0.359 (2.34)**
GCSE English grade			0.067 (6.35)***			0.050 (4.66)***
GCSE English grade missing flag			0.014 (0.10)			0.293 (2.68)***
Highest school qualification:						
One or two A levels		0.020 (0.93)	0.026 (1.18)		0.084 (3.39)***	0.083 (3.32)***
Three or more A levels		0.344 (21.44)***	0.254 (14.56)***		0.402 (22.74)***	0.341 (17.83)***
-2 (log_R – log_U)	183.5***	703.5***	3220.38***	234.2***	838.7***	957.76***
Observations		4883			4457	

Data: Youth Cohort Survey Sample with at least 5 Good GCSEs.

Notes: Dependent variable – value of one if in HE, zero otherwise, probit estimation. Base case: skilled background, white, father/ mother's education less than A level, attended a comprehensive and with no A levels.

All specifications also control for gender, ethnicity, parental education and school type. As for Table 5.

Absolute value of t statistics in parentheses * significant at 10 per cent; ** significant at 5 per cent; *** significant at 1 per cent.

**Appendix A: Probability of getting 5 or more good GCSEs
(Marginal effects)**

	1996	2000
Male	-0.058 (5.07)***	-0.060 (5.13)***
Parents' Socio-Economic Status:		
Professional, managerial & tech. occ.	0.056 (3.53)***	0.126 (7.98)***
Other non-manual occ.	0.023 (1.42)	0.085 (5.52)***
Semi-skilled occ.-manual	-0.052 (2.66)***	-0.067 (3.32)***
Unskilled occ.	-0.082 (2.36)**	-0.191 (5.39)***
Other	-0.067 (3.45)***	-0.098 (3.87)***
Ethnicity:		
Black	-0.191 (3.07)***	-0.078 (1.59)
Asian	-0.001 (0.05)	0.013 (0.57)
Other ethnicity	-0.140 (1.88)*	-0.031 (0.64)
Ethnicity missing flag	-0.100 (1.65)*	-0.073 (1.17)
Parental education:		
Father degree	0.141 (4.63)***	-0.018 (0.83)
Father at least one A level	0.171 (10.42)***	0.011 (0.54)
Father education missing flag	-0.057 (3.46)***	-0.114 (6.86)***
Mother degree	0.139 (3.45)***	0.082 (4.47)***
Mother at least one A level	0.132 (7.97)***	0.132 (7.06)***
Mother education missing flag	-0.047 (2.76)***	-0.043 (2.61)***
Type of school attended:		
Comprehensive age 16	-0.018 (1.46)	-0.034 (2.75)***
Grammar	0.337 (10.96)***	0.244 (8.29)***
Secondary modern	-0.153 (4.87)***	-0.099 (2.85)***
Independent	0.295 (12.84)***	0.176 (7.93)***
-2 (log_R – log_I)	1345.5***	1033.3***
Observations	7985	6268

Data: Youth Cohort Survey Full Sample.

Notes: Dependent variable – value of one if 5+ Good GCSEs, zero otherwise, probit estimation. Base case: skilled background, white, father/ mother's education less than A level, attended a comprehensive.

Absolute value of t statistics in parentheses * significant at 10 per cent; ** significant at 5 per cent; *** significant at 1 per cent.

Referees Appendix Table 1: Full results for Table 2 of main paper

	Dependent Variable: Ln (Number of HE students) in each postcode					
	OLS	OLS	Fixed Effects	Fixed Effects	Fixed Effects	Fixed Effects
Year:						
Year 1995/96	0.218 (15.26)***	-0.078 (0.50)	0.066 (1.10)	0.010 (0.16)	-0.027 (0.34)	0.071 (0.94)
Year 1996/97	0.371 (26.34)***	0.141 (0.91)	0.101 (1.71)*	0.019 (0.32)	-0.121 (1.57)	0.023 (0.31)
Year 1997/98	0.416 (29.67)***	0.270 (1.75)*	0.158 (2.67)***	0.076 (1.26)	0.008 (0.11)	0.147 (1.98)**
Year 1998/99	0.515 (36.73)***	0.425 (2.75)***	0.318 (5.41)***	0.276 (4.61)***	0.294 (3.83)***	0.365 (4.94)***
Year 1999/00	0.534 (38.11)***	0.314 (2.03)**	0.157 (2.67)***	0.122 (2.04)**	0.150 (1.95)*	0.208 (2.82)***
Year 2000/01	0.610 (43.50)***	0.525 (3.39)***	0.348 (5.91)***	0.311 (5.19)***	0.377 (4.90)***	0.436 (5.90)***
Year 2001/02	0.654 (46.65)***	0.599 (3.88)***	0.464 (7.89)***	0.418 (6.98)***	0.491 (6.39)***	0.564 (7.64)***
Population aged 18-24 from Census (in thousands)	0.773 (105.78)***	0.842 (125.92)***	-	-	-	-
Ln (Income)		1.333 (36.95)***	-	-	-	-
Interactions Ln (Income)*years:						
Ln (Income)*year 1995		0.096 (1.87)*	0.052 (2.67)***	0.061 (3.10)***	0.076 (2.71)***	0.033 (1.20)
Ln (Income)*year 1996		0.075 (1.48)	0.096 (4.95)***	0.108 (5.60)***	0.165 (5.99)***	0.095 (3.49)***
Ln (Income)*year 1997		0.047 (0.93)	0.099 (5.13)***	0.112 (5.78)***	0.139 (5.04)***	0.074 (2.74)***
Ln (Income)*year 1998		0.029 (0.57)	0.082 (4.27)***	0.089 (4.60)***	0.081 (2.95)***	0.051 (1.89)*
Ln (Income)*year 1999		0.071 (1.41)	0.142 (7.37)***	0.147 (7.62)***	0.136 (4.93)***	0.113 (4.18)***
Ln (Income)*year 2000		0.027 (0.53)	0.103 (5.37)***	0.109 (5.65)***	0.082 (2.97)***	0.058 (2.14)**
Ln (Income)*year 2001		0.017 (0.34)	0.080 (4.18)***	0.087 (4.53)***	0.058 (2.10)**	0.029 (1.06)
Interactions Population aged 18-24*years:						
Population 18-24*year 1995				0.051 (4.92)***	0.053 (4.97)***	
Population 18-24*year 1996				0.074 (7.25)***	0.080 (7.69)***	
Population 18-24*year 1997				0.074 (7.22)***	0.077 (7.35)***	
Population 18-24*year 1998				0.039 (3.81)***	0.038 (3.65)***	
Population 18-24*year 1999				0.031 (3.08)***	0.030 (2.90)***	
Population 18-24*year 2000				0.034 (3.33)***	0.031 (2.98)***	
Population 18-24*year 2001				0.042 (4.11)***	0.039 (3.72)***	
Ratio (Population Qualification level 4-5/Population aged older 24)						
Ratio Population aged older 24 Qualification level 4-5					-	-
Interactions Ratio (Population Qualification level 4-5/Population aged older 24)*years:						
Interaction Ratio Population aged older 24 Qualification level 4-5*year 1995					-0.046 (0.76)	0.088 (1.43)
Interaction Ratio Population aged older 24 Qualification level 4-5*year 1996					-0.172 (2.90)***	0.061 (1.00)
Interaction Ratio Population aged older 24 Qualification level 4-5*year 1997					-0.082 (1.39)	0.127 (2.09)**
Interaction Ratio Population aged older 24 Qualification level 4-5*year 1998					0.022 (0.38)	0.115 (1.89)*
Interaction Ratio Population aged older 24 Qualification level 4-5*year 1999					0.034 (0.57)	0.097 (1.61)

Interaction Ratio Population aged older 24 Qualification level 4-5*year 2000					0.081	0.152
					(1.37)	(2.50)**
Interaction Ratio (Population aged older 24 Qualification level 4-5)*year 2001					0.089	0.174
					(1.52)	(2.88)***
Population aged 10-15 from Census (in thousands)						-
						-
Interactions Population aged 10-15*years:						
Population 10-15*year 1995						0.064
						(3.78)***
Population 10-15*year 1996						0.124
						(7.46)***
Population 10-15*year 1997						0.107
						(6.48)***
Population 10-15*year 1998						0.042
						(2.52)**
Population 10-15*year 1999						0.024
						(1.44)
Population 10-15*year 2000						0.029
						(1.76)*
Population 10-15*year 2001						0.034
						(2.07)**
Constant	3.058	-1.046	3.465	3.465	3.465	3.465
	(281.04)***	(9.46)***	(997.21)***	(997.89)***	(998.12)***	(998.10)***
Observations	57743	57743	57743	57743	57743	57743
R-squared	0.20	0.34	0.40	0.40	0.40	0.40
Number of postcodes	-	-	7382	7382	7382	7382

Data: Higher Education Statistics Agency. Sample restricted to full and part-time students age 18-24 year olds domiciled in England or Wales enrolled in a first-degree course. See main paper for other sample restrictions.

**Referees Appendix Table 2: YCS Full Sample Probit Model (marginal effects) –
Specification as for Table 5 but without school type**

	1996		2000	
	Degree vs. other activities		Degree vs. other activities	
	Specification 1	Specification 2	Specification 1	Specification 2
Male	-0.009 (0.83)	0.000 (0.02)	-0.036 (2.89)***	-0.008 (0.63)
Parents' Socio-Economic Status:				
Professional, managerial & t.. occ.	0.056 (3.94)***	0.031 (2.15)**	0.138 (7.64)***	0.069 (3.64)***
Other non-manual occ.	0.029 (1.97)**	0.011 (0.72)	0.101 (5.62)***	0.048 (2.54)**
Semi-skilled occ.-manual	-0.045 (2.46)**	-0.040 (2.11)**	-0.026 (1.11)	-0.026 (1.03)
Unskilled occ.	-0.025 (0.76)	-0.005 (0.14)	-0.109 (2.82)***	-0.062 (1.41)
Other	-0.059 (3.30)***	-0.051 (2.74)***	-0.030 (1.03)	-0.033 (1.08)
Ethnicity:				
Black	-0.063 (1.13)	-0.039 (0.65)	-0.075 (1.43)	0.001 (0.02)
Asian	0.087 (3.83)***	0.108 (4.60)***	0.131 (5.07)***	0.163 (5.85)***
Other ethnicity	-0.026 (0.42)	-0.057 (0.92)	-0.066 (1.26)	-0.025 (0.45)
Ethnicity missing flag	-0.110 (1.86)*	-0.080 (1.25)	-0.009 (0.13)	0.018 (0.24)
Parental education:				
Father degree	0.135 (4.77)***	0.067 (2.37)**	0.022 (1.01)	0.031 (1.35)
Father at least one A level	0.145 (9.67)***	0.089 (5.83)***	0.107 (5.49)***	0.074 (3.61)***
Father education missing flag	-0.018 (1.12)	-0.002 (0.13)	-0.073 (4.06)***	-0.020 (1.00)
Mother degree	0.128 (3.52)***	0.074 (2.06)**	0.075 (3.75)***	0.042 (2.03)**
Mother at least one A level	0.117 (7.91)***	0.079 (5.25)***	0.093 (4.69)***	0.021 (1.02)
Mother education missing flag	-0.054 (3.39)***	-0.037 (2.26)**	-0.061 (3.29)***	-0.044 (2.20)**
Highest school qualification:				
One or two A levels		0.191 (10.68)***		0.227 (10.63)***
Three or more A levels		0.500 (37.11)***		0.520 (36.05)***
-2 (log_R – log_U)	535.5***	199.1***	546.9***	1964.6***
Observations		7985		6268

Data: Youth Cohort Survey Full Sample.

Notes: Dependent variable – value of one if 5+ Good GCSEs, zero otherwise, probit estimation. Base case: skilled background, white, father/ mother's education less than A level.

Absolute value of t statistics in parentheses * significant at 10 per cent; ** significant at 5 per cent; *** significant at 1 per cent.

**Referees Appendix Table 3: YCS Full Sample Probit Model (marginal effects) -
Specification as for Table 5 but including GCSE math & English grades**

	1996		2000	
	Degree vs. other activities	Degree vs. other activities	Degree vs. other activities	Degree vs. other activities
Male	-0.005 (0.47)	-0.004 (0.34)	-0.007 (0.55)	-0.001 (0.07)
Parents' Socio-Economic Status:				
Professional, managerial & t.. occ.	-0.007 (0.54)	-0.010 (0.74)	0.041 (2.22)**	0.027 (1.43)
Other non-manual occ.	-0.003 (0.20)	-0.005 (0.36)	0.033 (1.79)*	0.019 (1.03)
Semi-skilled occ.-manual	-0.014 (0.80)	-0.016 (0.87)	0.003 (0.10)	-0.007 (0.29)
Unskilled occ.	0.023 (0.69)	0.027 (0.78)	-0.035 (0.79)	-0.027 (0.57)
Other	-0.029 (1.64)*	-0.029 (1.65)*	0.000 (0.00)	-0.018 (0.54)
Ethnicity:				
Black	0.019 (0.32)	0.022 (0.36)	0.044 (0.75)	0.058 (0.98)
Asian	0.114 (4.85)***	0.117 (4.96)***	0.164 (5.94)***	0.178 (6.24)***
Other ethnicity	0.011 (0.18)	-0.016 (0.25)	-0.018 (0.34)	-0.003 (0.05)
Ethnicity missing flag	-0.059 (0.95)	-0.048 (0.75)	-0.008 (0.10)	0.012 (0.15)
Parental education:				
Father degree	-0.001 (0.04)	-0.009 (0.37)	0.002 (0.10)	0.014 (0.61)
Father at least one A level	0.004 (0.29)	0.001 (0.06)	0.048 (2.45)**	0.045 (2.22)**
Father education missing flag	0.009 (0.59)	0.012 (0.79)	-0.024 (1.26)	-0.004 (0.22)
Mother degree	0.008 (0.27)	0.001 (0.04)	0.013 (0.67)	0.015 (0.73)
Mother at least one A level	0.026 (1.97)**	0.022 (1.60)	-0.013 (0.68)	-0.026 (1.32)
Mother education missing flag	-0.028 (1.84)*	-0.024 (1.54)	-0.038 (1.94)*	-0.032 (1.63)
Type of school attended:				
Comprehensive age 16	-0.032 (2.90)***	-0.029 (2.57)**	-0.024 (1.61)	-0.021 (1.39)
Grammar	0.006 (0.27)	0.003 (0.16)	-0.063 (2.57)**	-0.073 (2.94)***
Secondary modern	-0.028 (0.96)	-0.022 (0.72)	-0.094 (2.30)**	-0.048 (1.10)
Independent	-0.003 (0.18)	0.017 (0.99)	-0.018 (0.84)	0.011 (0.50)
GCSE grades:				
GCSE maths grade	0.120 (24.01)***	0.100 (19.27)***	0.120 (18.64)***	0.081 (12.02)***
GCSE math grade missing flag	0.053 (0.84)	-0.001 (0.02)	0.065 (0.93)	-0.011 (0.16)
GCSE English grade	0.093 (15.58)***	0.072 (11.60)***	0.103 (13.37)***	0.065 (8.04)***
GCSE English grade missing flag	0.044 (0.55)	0.015 (0.19)	0.276 (4.08)***	0.211 (3.07)***
Highest school qualification:				
One or two A levels		0.074 (4.73)***		0.139 (6.54)***
Three or more A levels		0.228 (16.59)***		0.347 (20.73)***
-2 (log_R - log_U)	2943.1***	3220.4***	1950.8***	2396.7***
Observations	7985		6268	

Data: Youth Cohort Survey Full Sample.

Notes: Dependent variable – value of one if 5+ Good GCSEs, zero otherwise, probit estimation. Base case: skilled background, white, father/ mother's education less than A level, attended a comprehensive.

Absolute value of t statistics in parentheses * significant at 10 per cent; ** significant at 5 per cent; *** significant at 1 per cent.

Referees Appendix Table 4: The Determinants of HE Participation (Marginal effects) – Sample with 5+ Good GCSEs – full specification for Table 5

	1996			2000		
	Specif. 1	Specif. 2	Specif. 3	Specif. 1	Specif. 2	Specif. 3
Male	0.023 (1.56)	0.020 (1.28)	0.003 (0.19)	-0.024 (1.53)	-0.008 (0.47)	-0.010 (0.59)
Parents' Socio-Economic Status:						
Professional, managerial & t.. occ.	0.013 (0.69)	0.004 (0.20)	-0.008 (0.39)	0.088 (4.06)***	0.052 (2.28)**	0.032 (1.36)
Other non-manual occ.	0.016 (0.77)	0.004 (0.18)	-0.008 (0.38)	0.074 (3.39)***	0.042 (1.81)*	0.024 (1.04)
Semi-skilled occ.-manual	-0.003 (0.11)	-0.009 (0.32)	-0.004 (0.14)	0.011 (0.35)	-0.009 (0.28)	-0.011 (0.32)
Unskilled occ.	0.044 (0.84)	0.054 (1.00)	0.056 (1.00)	-0.010 (0.17)	0.003 (0.04)	0.009 (0.14)
Other	-0.035 (1.28)	-0.034 (1.19)	-0.029 (1.01)	0.045 (1.08)	0.014 (0.31)	0.011 (0.25)
Ethnicity:						
Black	-0.017 (0.17)	-0.038 (0.36)	-0.033 (0.31)	-0.048 (0.68)	0.021 (0.29)	0.060 (0.80)
Asian	0.094 (2.89)***	0.109 (3.28)***	0.114 (3.32)***	0.168 (5.27)***	0.193 (5.79)***	0.200 (5.97)***
Other ethnicity	-0.014 (0.15)	-0.054 (0.59)	-0.003 (0.03)	-0.128 (1.92)*	-0.101 (1.40)	-0.089 (1.22)
Ethnicity missing flag	-0.093 (0.95)	-0.054 (0.52)	-0.029 (0.27)	0.099 (0.98)	0.109 (1.02)	0.074 (0.68)
Parental education:						
Father degree	0.064 (1.82)*	0.028 (0.79)	-0.012 (0.34)	0.020 (0.77)	0.029 (1.04)	0.022 (0.80)
Father at least one A level	0.074 (3.86)***	0.048 (2.42)**	-0.000 (0.02)	0.110 (4.75)***	0.087 (3.57)***	0.071 (2.90)***
Father education missing flag	0.020 (0.87)	0.024 (0.98)	0.019 (0.77)	-0.028 (1.20)	0.006 (0.26)	0.013 (0.50)
Mother degree	0.066 (1.48)	0.033 (0.73)	-0.010 (0.21)	0.041 (1.75)*	0.029 (1.21)	0.019 (0.78)
Mother at least one A level	0.068 (3.62)***	0.048 (2.45)**	0.030 (1.48)	0.026 (1.13)	-0.013 (0.53)	-0.035 (1.43)
Mother education missing flag	-0.052 (2.21)**	-0.038 (1.58)	-0.030 (1.21)	-0.056 (2.32)**	-0.046 (1.78)*	-0.041 (1.59)
Type of school attended:						
Comprehensive age 16	-0.050 (2.94)***	-0.045 (2.56)**	-0.047 (2.58)***	-0.030 (1.64)	-0.028 (1.47)	-0.027 (1.37)
Grammar	0.122 (4.07)***	0.085 (2.71)***	0.003 (0.10)	0.045 (1.52)	-0.024 (0.78)	-0.082 (2.64)***
Secondary modern	-0.132 (2.71)***	-0.098 (1.93)*	-0.059 (1.13)	-0.180 (3.43)***	-0.077 (1.36)	-0.057 (0.99)
Independent	0.103 (4.45)***	0.107 (4.54)***	0.025 (1.00)	0.091 (3.60)***	0.078 (2.95)***	0.015 (0.55)
GCSE grades:						
GCSE maths grade			0.115 (13.10)***			0.074 (7.97)***
GCSE math grade missing flag			-0.138 (0.73)			0.359 (2.34)**
GCSE English grade			0.067 (6.35)***			0.050 (4.66)***
GCSE English grade missing flag			0.014 (0.10)			0.293 (2.68)***
Highest school qualification:						
One or two A levels		0.020 (0.93)	0.026 (1.18)		0.084 (3.39)***	0.083 (3.32)***
Three or more A levels		0.344 (21.44)***	0.254 (14.56)***		0.402 (22.74)***	0.341 (17.83)***
-2 (log_R – log_U)	183.5***	703.5***	3220.38***	234.2***	838.7***	957.76***
Observations		4883			4457	

Data: Youth Cohort Survey 5+ GCSE Sample.

Notes: Dependent variable – value of one if in HE, zero otherwise, probit estimation. Base case: skilled background, white, father/ mother's education less than A level, attended a comprehensive and with no A levels.

Absolute value of t statistics in parentheses * significant at 10 per cent; ** significant at 5 per cent; *** significant at 1 per cent.

