THE UNIVERSITY of York

Discussion Papers in Economics

No. 2008/23

The Socio-Economic Gap in University Drop Out

By

Nattavudh Powdthavee, University of London and University of York and
Anna Vignoles, University of London

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

The socio-economic gap in university drop out

Nattavudh Powdthavee
University of London and University of York

Anna Vignoles[1]
University of London

May 2008

ABSTRACT

It has been shown in the recent literature on widening participation that in England a disadvantaged pupil has as much a chance of attending a university as a more advantaged student, provided that s/he manages to reach a sufficient level of achievement at the secondary school level. This finding leads to an important conclusion of no genuine socio-economic gap in university participation once prior attainments have been taken into account. The current article investigates whether the same conclusion can be reached with respect to university drop-out. Using a combination of school and higher education administrative data sets, we are able to show that there is indeed a sizeable and statistically significant gap in the rate of withdrawal after the first year of university between the most advantaged and disadvantaged English students. This socio-economic gap in university drop-out remains even after allowing for their personal characteristics, prior achievement and institution choice. Our results thus suggest that the use of raw drop out rates in the English university 'league table' as one of the main indicators of university efficiency can be quite misleading given that the ranking of universities by drop out rate would change markedly if the prior attainment of students were taken fully into account.

Key words: Drop out rate; Higher Education; Prior achievement; Socio-economic gap
JEL: I2

[1] Centre for Economics of Education, Institute of Education. Corresponding author: a.vignoles@ioe.ac.uk

Introduction

As in the US and many other countries, the under representation of poorer students in college or university education has been an issue of great policy concern in the UK for many decades. Recent evidence for England indicates that the 20% most disadvantaged students are around 6 times less likely to participate in university compared to the 20% most advantaged pupils (HEFCE, 2005). Furthermore, the socio-economic gap in the university participation rate actually widened in the UK in the mid and late 1990s (Blanden and Machin, 2004; Machin and Vignoles, 2004; HEFCE, 2005). As in the US (Cunha et al. 2006), much of the root cause of this inequality is located earlier in the education system. Chowdry, Crawford, Dearden, Goodman and Vignoles (2008) have shown that in England, if a disadvantaged pupil does reach a sufficient level of achievement in secondary school, s/he has a similar chance of going on to university as a more advantaged student. However, in both the UK and the US, the lower participation rate of poorer students is not the only policy concern. To fully reap the rewards from a university education, poorer students need to complete their degrees. Drop out or non-completion has been seen as particularly problematic for students from disadvantaged backgrounds (Dearing, 1997; McGivney, 1996; HEFCE, 1999; Quinn, 2004)[2], and the extent to which drop out does indeed vary by socio-economic background is the focus of this paper.

Unlike in the US, the university sector in England has historically had low levels of student 'drop out' (Dearing, 1997; NAO, 2007). Recent data suggests that 91.6% of full time students starting university in 2004/05 continued into their second year and 78.1% are expected to complete their degree (NAO, 2007). However, as the sector has expanded and the rate of non-completion has risen (Johnes and McNabb, 2004), so policy attention has shifted to this issue, and non-completion rates are now part of a range of indicators used by government to measure university performance. Indeed university league tables are produced in UK newspapers, ranking universities on a number of metrics, including their 'drop out' rate. These league tables generally do not control for student characteristics, and therefore may give a misleading impression of the true institutional quality in terms of retaining students for universities with a

[2] The socio-economic gap in drop out from college in the US system is discussed extensively in Turner (2008).

larger number of deprived and lower achieving students. Ideally value added models, which control for students' prior achievement, are needed to assess whether institutions have particularly high or low drop out rates relative to their student intake. Such models have not been possible previously, due to limited data. However, in this paper we show how such models may be operationalised using administrative data.

Specifically in this paper we ask whether disadvantaged entrants to HE have a higher probability of 'dropping out', given their level of prior achievement. In other words, does disadvantage and poverty mean that although you can get in to HE you are then more likely to struggle when in HE and eventually 'drop out'. For example, we know that poorer students leave university with more debt and may be more risk averse in the first place (Pennell and West, 2005), so some have suggested that financial concerns may cause poorer students to drop out of university to a greater extent than their more affluent counterparts. This would imply that the focus of widening participation policy needs to be on facilitating degree completion by poorer students, rather than simply encouraging HE participation. Alternatively, is it simply the case that poorer students drop out of university not because they are poor but because they have lower levels of prior education achievement and are therefore less well prepared for HE? In other words it may be that poorer students are no more likely to drop out from HE than other more advantaged students with similar levels of (low) prior achievement.

To address these questions, we use administrative data on an entire cohort[3] of young people in England who potentially could enter university in 2003/4 (at age 18). These data are unique in that they include information on each pupil's personal characteristics (e.g. ethnicity, date of birth and indicators of their socio-economic background) and also provide a complete record of each child's educational achievement from age 11 onwards[4]. This is the first time that such longitudinal data has been available to study the important issue of drop out in HE in the UK context.

Previous literature

[3] We have data on all state school students in England who turned age 18 in 2003/4, i.e. more than half a million students.
[4] In the UK pupils take achievement tests or externally validated examinations at ages 11, 14, 16 and 18, hence we have a complete record of educational achievement from age 11 to age 19.

There is a large and growing literature on the role of family background, i.e. income and socio-economic status, in determining education outcomes, particularly university achievement (Blanden and Gregg, 2004; Carneiro and Heckman, 2002 and 2003; Gayle et al. 2002; Meghir and Palme, 2005; Haveman and Wolfe, 1995). Such studies have found family background to be an important determinant of educational achievement and have also suggested that the socio-economic gap in educational achievement emerges early (see CMPO, 2006 and Feinstein, 2003 for the UK; Cunha and Heckman, 2007 and Cunha et al., 2006 for the US). In fact Cunha et al. (2006) concludes that family background and specifically credit constraints play only a limited role in determining HE participation, conditional on achievement in secondary school, although some recent studies dispute this (Belley and Lochner, 2008).

Even if education inequality emerges early, in the US at least, the raw socio-economic gap widens substantially if one measures Bachelor degree completion as opposed to enrolment (Tuner, 2008). This raises the question as to whether the conditional drop out rate is higher for lower socio-economic group students, taking account of their prior education achievement. The literature on the relationship between socio-economic background and drop out from university is sparser, although the US evidence reports differential drop out by family income (see Corrigan, 2003, Haveman and Wilson, 2005 and related issues in Bound, Lovenheim and Turner, 2008). In the English context, Johnes & McNabb (2004) analyzed students entering and leaving the 'old' (pre-1992[5]) universities and distinguished between 'voluntary' drop out and 'involuntary' drop out i.e. failure. Jones & McNabb found that students from a lower socio-economic background were more likely to drop out voluntarily. Smith & Naylor (2001) used the same data to examine completion and non-completion. Using a binomial regression analysis of the probability that an individual withdraws from university for whatever reason, the authors found the risk of dropping out to be extremely high amongst students from lower social class backgrounds and living in high unemployment rate areas. More recently, Bennett (2003) showed self-declared financial hardship to be the most powerful predictor of a student's decision to withdraw from their degree course in the Business Studies department in a 'new'

[5] Prior to 1992, English higher education was divided into universities (higher status more academically oriented universities) and polytechnics (more vocationally oriented higher education institutions). In 1992 polytechnics became universities.

university in Greater London. Using data collected by the author himself, Bennett estimated a confirmatory factor analysis of the probability of non-completion among business students. He found self-declared financial hardship to be one of the strongest predictors of the individual quit decision. Other important factors also included low self-esteem and academic performance at the university. Whilst these studies were able to control to some extent for a student's entry qualifications, they did not have rich data on students' prior achievement, and only considered a sub-set of UK universities.

This paper therefore estimates models of student drop our or non-completion using all universities in England. Before we do this, it is worth considering why we might be concerned with student drop out per se, as opposed to differential drop out by socio-economic background. Firstly, there may be economic costs associated with non-completion; there may have been a waste of resources if a student starts but does not complete a course (Yorke, 1998). Another potential concern is the sense of failure that a student may feel after dropping out of university and the impact on his or her earnings. At the same time, there is recognition that for many students, progression through university is not linear. This is particularly true of mature students (McGivney, 1996). Labelling (temporary) withdrawal as academic failure or wastage would seem inappropriate: just because students withdraw from their studies does not mean that they have not received any benefit from university (Johnes and Taylor, 1989). This is not merely a semantic debate: universities in England now face clear incentives to encourage student completion in the 'normal' time and non-completion whatever the cause is penalized financially[6]. If indeed poorer students are more likely to drop out than their more advantaged counterparts for a given level of prior achievement, this may lead to a tension between the widening participation agenda and the desire by universities not to incur penalties from high levels of student withdrawal (Palmer, 2001).

Some economists have also made the argument that drop out from HE is efficient: weaker students who would not benefit from completing their degrees rightly drop out. Manski (1989) for example, argues that lower dropout rates would not necessarily

[6] Public funding received by each university in England is potentially affected by their performance. Non-completion is one measure of university performance used by the authorities.

make society better off. He suggests the decision to enrol is a decision to initiate an experiment, a possible outcome of which is dropout (see also Hartog et al 1989; Oosterbeek 1989; and Altonji 1993). Thus enrolment in HE incurs a risk for all students, namely the risk that they may have to drop out for whatever reason. Poorer students may face higher levels of this risk. For example, they may be more likely to fail to reach the level of educational achievement required or make their decisions about choice of institution and subject of study on the basis of poorer quality information. This higher level of risk may partially explain lower participation rates by poorer students. Even if poorer students face the same risk of drop out as their advantaged peers, if they are more risk adverse (Callender, 2003), then this too would at least partially explain their lower enrolment rate.

Data

We use linked administrative data sets to carry out the analysis: namely, the English National Pupil Database (NPD) / Pupil Level Annual School Census (PLASC) and individual student records maintained by the Higher Education Statistics Agency (HESA). The school administrative data (NPD/PLASC)[7] contain each pupil's record of their primary and secondary schooling. In our data set specifically, we have information on each pupil's educational attainment from age 11 to 18, as well as their personal characteristics, such as date of birth, ethnicity, home postcode, entitlement to free school meals[8] and whether English is an additional language in their home. The university records (HESA data) contain information on the degree subject, institution and other details of each student's university education for all students studying for a first degree at UK universities. With these two sources of data linked together[9], we have longitudinal data on a cohort of students from age 11 through to potential HE participation at age 18 in 2004/5. For the purposes of this paper, we consider only HE participants and have a sample size of 128,423 observations from 161 HE institutions.

[7] These data are maintained by the Department for Children, Schools and Families.

[8] This can be thought of as a proxy for very low family income. Pupils are eligible for Free School Meals (FSM) if their parents receive Income Support, income-based Jobseeker's Allowance, or Child Tax Credits, with a gross household income of less than £14,495 (in 2007–08 prices).

[9] School administrative records are fuzzy matched to higher education administrative records, using a variety of variables including name, date of birth and postcode. The matching process was carried out by the Department for Children, Schools and Families.

The dependent variable of interest is simply whether or not the pupil continued from one year to the next, i.e. continued in the same university from 04/05 to 05/06. Around 6% of pupils failed to progress from one year to the next, indicating they dropped out from their institution (voluntarily or involuntarily) or decided to move to another institution in the following year. As our cohort only potentially entered HE the previous year, we are essentially measuring drop out after year 1 of 3 years of study[10].

A key feature of the data we use is that they include test score information on pupils from age 11 onwards. The test score information comes from age 11 (Key Stage 2) and age 14 (Key Stage 3) tests. These are national achievement tests sat by all children in state schools in England in English, Mathematics and Science. The tests are externally validated i.e. they are marked by individuals outside of the child's school. We take the actual marks obtained by the child in these tests and average them across the three subject areas – English, Mathematics and Science. We then generate quintiles from this continuous average score to better identify any non-linearities in the effects of these measures of prior achievement (see Chowdry et al. 2008 for full details of the methodology used). The test data are supplemented by the results from public examinations taken by most students at age 16, namely General Certificates of Secondary Education (GCSEs), and for some students, Advanced levels (A-levels) at age 18. For GCSE, we use the capped total point score: this gives the total number of points accumulated from the student's eight highest GCSE grades.[11] At 18, we use the total (uncapped) point score. As with age 11 and age 14 test scores, we divide the population into five evenly sized quintile groups ranked according to their score at GCSE and A level or equivalent[12] to capture attainment at these levels. All in all these data contain the richest possible information on students' prior achievement to better enable us to identify the distinct role of academic preparation and socio-economic background in dropping out of university.

[10] The vast majority of UK Bachelors degree are 3 years of study.

[11] We use a capped total point score to avoid conflating the quantity of GCSEs taken and the grades received, as students may take a varying number of qualifications. For example, receiving 10 Grade D GCSEs would be equivalent (in terms of total points scored) to receiving 8 Grade C GCSEs (using the old tariff system), while we may not believe these are equivalent in terms of ability.

[12] For students taking vocational qualifications at age 18 instead of A-levels, we have their point score using the official equivalencies between vocational and academic qualifications.

Based on the person's university, we also linked in an institution-level indicator of the university's research quality from the 2001 Research Assessment Exercise (RAE)[13]. We then combine this indicator of the quality of each institution's research, with an indicator of whether or not the institution is a member of the Russell Group university, a self-defined elite grouping of English universities. We then define a high status institution as being all 20 of the research-intensive Russell Group institutions, plus any UK university with an average 2001 RAE rating that exceeds the lowest average RAE found among Russell Group universities (see Chowdry et al. 2008 for further details and a list of institutions). In summary, we create a binary indicator of whether an institution is an elite institution or otherwise.

The data do however, have a number of limitations. Firstly, the indicators for students' family background are somewhat crude. We have an indicator of whether or not a student was eligible for free school meals (FSM) in secondary school. Around 5% of students entering HE were eligible for free school meals in secondary school. Additionally we have each pupil's postcode[14] and can link in information on the characteristics of the pupil's neighborhood, particularly measures of socio-economic deprivation such as the unemployment rate. Each pupil's socio-economic background is then represented by his or her score on an index constructed by combining together (using principal component analysis[15]) three different measures of deprivation: the pupil's eligibility for Free School Meals (recorded at age 16), their Index of Multiple Deprivation (IMD) score (derived from Census data on the characteristics of individuals living in their neighborhood[16]) and their Income Deprivation Affecting Children Index (IDACI) score, again constructed on the basis of Census data on individuals living in their neighborhood[17]. The population is split into five quintiles on the basis of this index. Whilst these measures of family deprivation are not ideal (family income would be preferable, for example), taken together they provide a clear indicator of the deprivation of any given pupil.

[13] The RAE is a quality assessment exercise to assess research quality across the HE sector in England and Wales. Quality assessment is based on peer review.
[14] Geographic identifier akin to a zip code.
[15] This method takes into account the different scales of the contributing variables.
[16] This is available at Super Output Area (SOA) level (comprising approximately 700 households), and makes use of information from seven different domains: income; employment; health and disability; education, skills and training; barriers to housing and services; living environment; and crime.
[17] IDACI is an additional supplementary element to the Index of Multiple Deprivation.

Another limitation of the data is that we only consider young HE participants i.e. those participating at age 18. The drop out behaviour of older HE participants may differ so our results are not necessarily generalisable to older students. Finally, we only have data on state school pupils. A significant minority of students in England attend private schools prior to entering HE (just under 7% at age 16 in our data). If these more advantaged students were included in our sample and if they have very low drop out rates, then our estimates of the socio-economic gap in HE drop out may well be lower bounds.

Table 1: Descriptive Statistics of Selected Variables

Variables	Enter HE at 18			Enter HE at 19
	All	Continued	Withdrawal	All
Male	0.432	0.433	0.429	0.477
	(0.495)	(0.495)	(0.495)	(0.499)
White	0.726	0.722	0.788	0.684
	(0.445)	(0.447)	(0.408)	(0.464)
Black	0.025	0.026	0.025	0.039
	(0.158)	(0.158)	(0.157)	(0.195)
Asian	0.096	0.098	0.056	0.102
	(0.294)	(0.297)	(0.231)	(0.303)
GCSE (capped): in quintile	4.336	4.377	3.684	3.940
	(0.823)	(0.789)	(1.052)	(1.067)
A-level or equiv points	290.094	297.441	171.146	193.977
	(143.767)	(140.659)	(141.275)	(166.417)
Multiple deprivation index: in quintile	2.411	2.397	2.631	2.585
	(1.301)	(1.297)	(1.347)	(1.374)
OA education index: in quintile	3.606	3.623	3.333	3.632
	(1.269)	(1.263)	(1.338)	(1.298)
University status index	0.374	0.387	0.171	0.279
	(0.484)	(0.487)	(0.377)	(0.448)
N	128423	120951	7472	56001

In Table 1, we show some descriptive statistics for our sample, including a comparison with students who entered HE one year later at age 19. Whilst we cannot include those who enter HE at 19 in 2005/6 in our analysis (because we do not have data on these students the following year to measure their propensity to withdraw), it is useful to see how the characteristics of later entrants compare with our sample. In general, those entering later appear to be more likely to be non-white, have lower achievement, are somewhat more deprived and have a lower probability of attending a high status institution. Such differences are not statistically significant however.

In our sample of those who entered HE at age 18, ‘drop outs’ are slightly (but not significantly) more likely to be white. Those who do not continue beyond their first year of study are lower achievers at GCSE and A level. They are slightly more deprived on our measure of deprivation and are much less likely to attend a high status university.

Econometric model

We define a model of non-continuation as follows:

$$\Pr(Y = 1 \mid X = x) = \Phi(x'\beta),$$
$$\Pr(Y = 1 \mid X = x)_{ij} = X_{ij}^{'}\beta + HE_{j}^{'}\gamma + \lambda_j + \varepsilon_{ij},$$

where Y takes a value of 1 if individual i from university j withdraws from the institution in the following academic year. We cannot distinguish between voluntary and involuntary withdrawal.

$X_{ij}^{'}$ is a vector of student personal characteristics taken largely from the secondary school administrative data set. This includes date of birth, gender, ethnicity, disability status, and English as an Additional Language. Our explanatory variables of interest are firstly our measures of the student's socio-economic status, in particular his or her quintile of socio-economic deprivation, as described in the previous section. Our second set of explanatory variables of interest are our comprehensive measures of prior educational attainment, namely test scores at age 11, 14, 16 (GCSE) and age 18 (A-level or equivalent).

$HE^{'}_{j}$ is a vector of variables describing the nature of the individual's university and degree subject. For some models we include detailed information on university quality, based on the institution-level Research Assessment Exercise (RAE) scores for 2001 and a dummy variable representing whether the institution is a Russell group university[18]. When we investigate interactions, we simplify this information into a simple binary variable indicating whether the university attended is a high status institution, as described earlier in the data section.

There are a number of estimation issues. We are attempting to determine whether or not poorer students are more likely to drop out of university, allowing for their prior educational achievement. Whilst we are confident that we can control fully for each

[18] Or an equivalent version if it is a Scottish institution.

pupil's prior achievement, which previous research has been unable to do, we recognize that there may be unobserved factors that determine whether or not a pupil drops out of university and that these may be correlated with their socio-economic status. For example, poorer students may tend to enroll in certain types of universities. There may be unobserved features of a student's university that then determine their decision to withdraw. Whilst we control for degree subject and institution quality in a relatively comprehensive way, there are many other features of universities that we are unable to allow for, such as pastoral care, advice and guidance given etc. To address this we therefore estimate fixed effect models, with a fixed effect for each university, which allows for any unobserved differences in the non-continuation rate across different HEIs. The variable λ_j represents institutional fixed effects (dummy variables for each institution bar the base case). The parameter ε_{ij} denotes the error term. The model is then estimated using a probit model, and marginal effects are reported in all tables, standard errors are clustered by university. We also allow for the potentially heterogeneous effects from socio-economic status by exploring various interactions as discussed in the section below.

Results

Table 2 below estimates the model described above, where the dependent variable takes a value of one if the person continued beyond their first year of study in the same university and zero otherwise. We then add sequentially various explanatory variables as controls. Column 1 presents the raw socio-economic gradient in 'drop out': it controls for the student's socio-economic background and ethnicity only. The probability of 'dropping out' reduces monotonically and statistically significantly as we move down the deprivation quintiles. Students from the lowest deprivation quintile i.e. the least deprived students were 3 percentage points less likely to withdraw after their first year of university as compared to the most deprived students in the highest deprivation quintile. Another measure of the student's socio-economic status is the education level of their neighborhood. There is also a clear negative relationship between the probability of withdrawal and neighborhood education levels. Students from the most educated neighborhood (highest OA education quintile) were almost 2 percentage-points less likely to withdraw compared to those

from the lowest OA education quintile. In the context of low overall drop our rates (6% in our sample), these socio-economic and ethnic differences are very large.

There also appears to be an ethnicity dimension to the problem of 'drop out'. Almost all groups of ethnic minority students are significantly less likely to dropout compared to white students. For example, Black African, Bangladeshi and other Asian students were approximately 3 percentage-points less likely to withdraw than Whites. Chinese students were nearly 5 percentage-points less likely to dropout compared to students of white ethnic background.

Column 2 then adds a variety of student personal characteristics, including gender, date of birth, and disability status. These variables are not statistically significant and make little difference to the socio-economic and ethnic variable coefficients.

Column 3 then adds controls for a number of university characteristics, including measures of university status and subject of degree. There are notable drops in the coefficients on the socio-economic variables, suggesting that some of the apparently higher withdrawal rate of poorer students is attributable to the types of degree subjects they are studying and the institution type they enroll in. For example, students from the lowest deprivation quintile are now just over 2 percentage-points less likely to withdrawal (compared to almost 3 percentage-points from column 1). It is not clear however, whether this would be a preferred specification. Since students' choice of degree and institution is itself related to their socio-economic background, this specification removes some of the socio-economic effect we are trying to measure. That said, it is useful to identify the routes by which the large raw differential in drop out rates across different socio-economic groups occurs.

In Column 3, the size of the ethnic variable coefficients also dropped somewhat when controls for the type of university and degree subject are included. For instance, Indian students are now 2.9 percentage-points less likely to dropout once HE characteristics are controlled for (compared to 3.6 percentage-points in the first column). The model also suggests that students enrolled in higher status universities were 2.6 percentage-points more likely to progress to the next year at the same university. The 'drop out' rate varies substantially by degree subject. This is

illustrated by the coefficients on the art and agricultural subject variables. Art and agricultural students were 11 percentage-points and 18 percentage-points respectively more likely to withdrawal from their course compared to medical students.

Column 4 then adds in our measures of prior attainment, namely achievement at ages 11, 14, 16 and 18. Further reductions in the coefficients on the socio-economic variables can be observed, though the socio-economic coefficients remain statistically significant. So the least deprived students (from the lowest deprivation quintile) are now one percentage point less likely to withdraw than their more deprived counterparts. This result implies that not all the socio-economic gradient in 'drop out' is attributable to the poorer prior academic achievement of disadvantaged pupils. In the light of previous comments about sorting across HE institutions by socio-economic background, we estimated this full specification, including prior achievement but without any university variables. This model generated very similar results to the model that included university characteristics, i.e. both indicated that students from the most deprived backgrounds are around 1 percentage point less likely to drop out than their wealthier peers. This suggests that sorting across universities and degree subjects is not as important as prior achievement in determining student drop out.

Controlling for prior attainment, students from Black Caribbean background are now 2 percentage-points more likely to remain studying in their institution beyond the first year of their degree. For Chinese students, on the other hand, the inclusion of prior attainment reduces the differential between their drop out rate and those of whites. Chinese students are now only 3 percentage-points less likely to withdraw as compared to whites (previously in the specification without prior attainment they were 5 percentage points less likely).

The inclusion of prior attainment measures in Column 4 also causes the coefficients on the degree subject variables to become insignificant, suggesting much of the apparent difference in the drop out rate across subjects is actually attributable to differential prior achievement of students (and the fact that students of different levels of prior achievement sort themselves into different degree subjects). In terms of the prior attainment variables themselves, generally only measures of prior achievement

at ages 16 and 18 are significant. For example, students in the highest GCSE quintile were approximately 6 percentage-points less likely to withdraw compared to those in the lowest GCSE quintile. Clearly these measures of prior achievement are of more importance, quantitatively speaking, than the student's socio-economic background.

As discussed earlier, we also want to control for the unobserved characteristics of universities that may impact on student 'drop out'. We do this by including institution fixed effects in Column 5. The inclusion of these fixed effects does not change the coefficients on the socio-economic variables in a major way: i.e. there remains a socio-economic gradient in university 'drop out' even after controlling for prior attainment and institution fixed effects. However, the inclusion of institution fixed effects does alter the coefficient on the university status variable. Without fixed effects, the model suggests that students from high status universities are no more likely to 'drop out' than students from other institutions. After including fixed effects however, students from high quality institution are now 2 percentage-points *more* likely to withdrawal. The inclusion of institution fixed effects does not however, alter the finding that there are no significant differences in 'drop out' across subject areas, once we control for prior achievement.

Table 2: The socio-economic gradient in HE non-continuation

Variables	Socio-economic	Plus personal characteristics	Plus HE characteristics	Plus prior attainment	Plus Institution fixed effects
4th deprivation quintile	-0.0125	-0.0124	-0.0106	-0.0060	-0.0063
	[0.0041]**	[0.0041]**	[0.0025]**	[0.0022]**	[0.0019]**
3rd deprivation quintile	-0.0177	-0.0175	-0.0146	-0.0067	-0.0066
	[0.0048]**	[0.0047]**	[0.0030]**	[0.0025]**	[0.0019]**
2nd deprivation quintile	-0.0238	-0.0234	-0.0200	-0.0100	-0.0093
	[0.0052]**	[0.0051]**	[0.0035]**	[0.0028]**	[0.0022]**
Lowest deprivation quintile	-0.0290	-0.0285	-0.0240	-0.0119	-0.0115
	[0.0058]**	[0.0057]**	[0.0040]**	[0.0030]**	[0.0022]**
2nd OA education quintile	-0.0018	-0.0015	-0.0012	0.0008	0.0004
	[0.0030]	[0.0029]	[0.0026]	[0.0024]	[0.0023]
3rd OA education quintile	-0.0090	-0.0086	-0.0066	-0.0028	-0.0034
	[0.0030]**	[0.0029]**	[0.0024]**	[0.0022]	[0.0019]
4th OA education quintile	-0.0163	-0.0160	-0.0114	-0.0065	-0.0067
	[0.0035]**	[0.0034]**	[0.0026]**	[0.0024]**	[0.0020]**
Highest OA education quintile	-0.0186	-0.0183	-0.0138	-0.0071	-0.0082
	[0.0044]**	[0.0043]**	[0.0032]**	[0.0027]**	[0.0022]**
Other White	-0.0065	-0.0063	-0.0048	-0.0075	-0.0083
	[0.0052]	[0.0052]	[0.0045]	[0.0037]*	[0.0034]*
Black African	-0.0266	-0.0258	-0.0209	-0.0255	-0.0261
	[0.0062]**	[0.0063]**	[0.0052]**	[0.0031]**	[0.0023]**
Black Caribbean	0.0049	0.0050	-0.0085	-0.0201	-0.0212
	[0.0144]	[0.0143]	[0.0067]	[0.0034]**	[0.0028]**
Other Black	-0.0225	-0.0220	-0.0210	-0.0255	-0.0253
	[0.0072]**	[0.0072]**	[0.0061]**	[0.0040]**	[0.0036]**
Indian	-0.0366	-0.0360	-0.0288	-0.0283	-0.0261
	[0.0043]**	[0.0042]**	[0.0023]**	[0.0019]**	[0.0014]**
Pakistani	-0.0213	-0.0209	-0.0164	-0.0213	-0.0198
	[0.0046]**	[0.0046]**	[0.0036]**	[0.0025]**	[0.0021]**
Bangladeshi	-0.0259	-0.0252	-0.0185	-0.0211	-0.0190
	[0.0063]**	[0.0062]**	[0.0045]**	[0.0036]**	[0.0031]**
Chinese	-0.0456	-0.0453	-0.0366	-0.0330	-0.0321
	[0.0052]**	[0.0052]**	[0.0034]**	[0.0025]**	[0.0020]**
Other Asian	-0.0292	-0.0289	-0.0184	-0.0159	-0.0157
	[0.0063]**	[0.0062]**	[0.0063]**	[0.0056]**	[0.0053]**
Mixed ethnicity	-0.0281	-0.0277	-0.0213	-0.0182	-0.0178
	[0.0057]**	[0.0057]**	[0.0046]**	[0.0043]**	[0.0040]**
Other ethnicity	-0.0140	-0.0140	-0.0140	-0.0192	-0.0199
	[0.0065]*	[0.0065]*	[0.0044]**	[0.0029]**	[0.0024]**

Unknown ethnicity	0.0006	0.0000	0.0019	-0.0012	-0.0012
	[0.0039]	[0.0038]	[0.0031]	[0.0026]	[0.0026]
Date of birth		0.0000	0.0000	0.0000	0.0000
		[0.0000]	[0.0000]*	[0.0000]**	[0.0000]**
Male		0.0006	0.0023	-0.0039	-0.0029
		[0.0026]	[0.0018]	[0.0015]**	[0.0014]*
Disabled		0.0146	0.0022	-0.0067	-0.0091
		[0.0080]	[0.0056]	[0.0033]*	[0.0021]**
Good university			-0.0259	0.0028	0.0243
			[0.0068]**	[0.0069]	[0.0033]**
Subject: Allied to medicine			0.0659	0.0006	-0.0005
			[0.0217]**	[0.0097]	[0.0091]
Subject: Biological science			0.0351	0.0018	-0.0011
			[0.0165]*	[0.0100]	[0.0090]
Subject: Veterinary science			0.0918	0.0162	-0.0024
			[0.0432]*	[0.0185]	[0.0123]
Subject: Agriculture			0.1846	0.0525	0.0143
			[0.0663]**	[0.0276]	[0.0143]
Subject: Physics			0.0254	-0.0010	-0.0017
			[0.0132]	[0.0084]	[0.0079]
Subject: Mathematics			0.0148	-0.0009	0.0006
			[0.0117]	[0.0081]	[0.0079]
Subject: Computer science			0.0670	0.0099	0.0052
			[0.0226]**	[0.0116]	[0.0101]
Subject: Engineer & Technology			0.0622	0.0108	0.0087
			[0.0222]**	[0.0118]	[0.0109]
Subject: Architecture			0.0481	0.0049	0.0007
			[0.0212]*	[0.0113]	[0.0096]
Subject: Social studies			0.0445	0.0072	0.0036
			[0.0175]*	[0.0104]	[0.0092]
Subject: Law			0.0211	0.0018	-0.0022
			[0.0141]	[0.0097]	[0.0085]
Subject: Business			0.0634	0.0144	0.0068
			[0.0207]**	[0.0115]	[0.0098]
Subject: Mass communication			0.0790	0.0229	0.0068
			[0.0401]*	[0.0194]	[0.0105]
Subject: Languages			0.0324	0.0068	0.0042
			[0.0146]*	[0.0099]	[0.0092]
Subject: History			0.0277	-0.0007	-0.0010
			[0.0171]	[0.0107]	[0.0102]
Subject: Arts			0.1121	0.0381	0.0200
			[0.0372]**	[0.0192]*	[0.0123]
Subject: Education			0.0575	0.0049	-0.0022

			[0.0221]**	[0.0113]	[0.0090]
Subject: Combined			0.0838	0.0169	0.0156
			[0.0268]**	[0.0139]	[0.0131]
2nd KS2 quintile				-0.0004	-0.0001
				[0.0031]	[0.0030]
3rd KS2 quintile				0.0004	0.0008
				[0.0035]	[0.0033]
4th KS2 quintile				0.0055	0.0055
				[0.0041]	[0.0037]
Highest KS2 quintile				0.0042	0.0047
				[0.0039]	[0.0036]
2nd KS3 quintile				-0.0022	-0.0002
				[0.0037]	[0.0036]
3rd KS3 quintile				-0.0053	-0.0034
				[0.0038]	[0.0036]
4th KS3 quintile				-0.0026	-0.0005
				[0.0042]	[0.0039]
Highest KS5 quintile				-0.0012	0.0014
				[0.0045]	[0.0041]
2nd capped KS4 quintile				-0.0099	-0.0100
				[0.0054]	[0.0052]
3rd capped KS4 quintile				-0.0273	-0.0251
				[0.0038]**	[0.0039]**
4th capped KS4 quintile				-0.0381	-0.0354
				[0.0048]**	[0.0048]**
Highest capped KS4 quintile				-0.0599	-0.0546
				[0.0078]**	[0.0078]**
KS5 points				-0.0002	-0.0002
				[0.0000]**	[0.0000]**
Level 3 at 18				-0.0256	-0.0218
				[0.0046]**	[0.0040]**
Institution fixed effects	No	No	No	No	Yes
Observations	128423	128423	128423	128423	127916
Log-likelihood	-28106.8481	-28021.337	-25585.6946	-24168.5407	-23605.4805
psuedo_r2	0.0139	0.0169	0.1023	0.1520	0.1709

One motivation for this paper was to explore whether universities themselves have different rates of ‘drop out’, once we control for the prior achievement of their students. This is important given the policy importance of university ‘league tables’ in the UK and the notion of ‘drop out’ rates being one metric by which university quality is judged. The model above, in column 5 allows for unobserved characteristics of

universities that remain constant and influence their average ‘drop out’ rate. Table 3 then investigates the magnitude of these institutional fixed effects further. It shows the size and significance of 10 randomly chosen fixed effects from the model estimated in Table 2. In column 1 of Table 3 the fixed effects for specific institutions are shown when the model includes no controls. Column 2 then shows the magnitude of these fixed effects when the model is estimated with full controls, including prior achievement of students (i.e. the specification in column 5 of Table 2). Table 3 therefore illustrates the difference that controlling fully for student characteristics and particularly prior achievement can make. For example, column 1 suggests that, in a drop out equation with institution dummies and no other controls, students attending University A are significantly less likely to ‘drop out’ i.e. 1.9 percentage points less likely, than the base institution. Once we control for student characteristics and prior achievement however, this is reduced to just 0.5 percentage points less. University D looks like it has a higher ‘drop out’ rate than the base case (by 4 percentage points). Yet once we control for student characteristics and prior achievement students from University D face similar risk of dropping out as students from the base case. We can conclude that any assessment of the continuation rates of universities needs to take full account of student characteristics and prior achievement if the results are not to be misleading.

Table 3: Randomly selected institutional dummies

Variables	No controls	Full controls	T-statistics (Test of equality)
Institution A	-0.0192	-0.0056	-1.422
	[0.0095]	[0.0011]**	
Institution B	-0.0379	-0.0316	-1.102
	[0.0057]**	[0.0004]**	
Institution C	-0.0435	-0.0332	-6.762**
	[0.0014]**	[0.0006]**	
Institution D	0.0360	-0.0015	0.410
	[0.0913]	[0.0039]	
Institution E	-0.0400	-0.0338	-1.106
	[0.0056]*	[0.0003]**	
Institution F	-0.0292	-0.0214	-2.336*
	[0.0033]**	[0.0005]**	
Institution G	-0.0320	-0.0234	-2.502*
	[0.0034]**	[0.0005]**	
Institution H	-0.0372	-0.0290	-4.246**
	[0.0018]**	[0.0007]**	
Institution I	-0.0338	-0.0226	-3.782**
	[0.0029]**	[0.0006]**	
Institution J	-0.0192	-0.0056	-1.73
	[0.0078]	[0.0011]**	
Observations	127916	127916	
Log-likelihood	-25633.171	-23605.481	
Pseudo R-squared	0.0997	0.1709	

Fixed effects in the second column taken from the model estimated in Table 2. No clustering of HEI in the first column with no other controls.

Results by gender and for higher status (Russell Group) universities and Black students are reported in the appendix. Whilst the coefficients on some variables do vary across these different groups of students and institutions, the results suggest that the relationship between socio-economic background and student 'drop out' is quite similar by gender and institution type (with a marginally smaller socio-economic gradient for students at higher status institutions – but not significantly so). For Black students by contrast, there appears to be very little socio-economic gradient in 'drop out' once you control fully for prior achievement.

As another robustness check, we also explored interactions between socio-economic variables and the type of university attended. Here, we used the binary variable as to whether the university was a higher status institution (see data section for a

description of what constitutes a higher status institution). We found that the interactions between attending a higher status university and socio-economic background were not significant (results in the Appendix). Thus it appears not to be the case that poorer students fare better at certain types of HE institutions than others, in terms of their likelihood of 'dropping out'. Some of the interaction effects between ethnicity and university status were significant however. Specifically the coefficients on the interactions between variables indicating the student is of Pakistani and other Asian ethnicity and university status. These interactions generally cancelled out the main effects of ethnicity dummies.

Conclusions

Our results suggest that, in the UK, as in the US, there is a significant gap in the drop out rate between advantaged and disadvantaged pupils. In the context, of a relatively low aggregate rate of first year drop out (6%) from English universities, the 20% most deprived pupils in England are around 3 percentage points more likely to drop out than the most advantaged 20%. Much of this gap disappears once we allow for students' prior achievement, suggesting that some of the apparent difference in first year drop out rates between richer and poorer students is actually attributable to differences in their academic preparation for HE and/or their ability, as measured by earlier measures of educational achievement. That said, there remains a 1 percentage point difference in the drop out rate of the most advantaged and disadvantaged English students, even allowing for their personal characteristics, prior achievement and institution choice. In the context of a low overall drop out rate in the UK, this 1 percentage point difference in drop out rates between richer and poorer students, even after controlling for prior achievement, is arguably sizeable.

We also found that raw indicators of the drop out rate of English universities could be quite misleading, if one's purpose is to use such measures as indicators of university efficiency. For instance, the ranking of universities by drop out rate would change markedly if the prior achievement of students were taken fully into account. In policy terms, this suggests that if we are to use drop out rates as measures of institution performance, we must be careful to apply a value added model to the data first.

Undoubtedly, the English system does not yet have the high drop out rates of the US system. However, the drop out rate has risen in recent years in the UK and the evidence here suggests that we should be alert to the fact that this will tend to widen the socio-economic gap in degree completion, since poorer students drop out to a greater extent even after allowing for their prior achievement.

Reference

Allen, D. (1999). Desire to finish college: an empirical link between motivation and persistence. *Research in Higher Education*, 40(4), 461-485.

Altonji, J.G. (1993). The demand for and return to education when education outcomes are uncertain. *Journal of Labor Economics* 11 1, pp. 48–83.

Belley and Lochner (2008) *Human Capital*

Bennett, R. (2003). Determinants of undergraduate student drop out rates in a university business studies department. Journal of Further and Higher Education, 27(2), 123-141.

Blanden, J. and P. Gregg (2004), *Family Income and Educational Attainment: A Review of Approaches and Evidence for Britain*, Oxford Review of Economic Policy, Vol. 20, pp. 245-263

Blanden, J. and S. Machin (2004), *Educational Inequality and the Expansion of UK Higher Education*, Scottish Journal of Political Economy, Special Issue on the Economics of Education, Vol. 51, pp. 230-49

Bound, J. and Turner, S. (2007) *Cohort Crowding: How Resources Affect Collegiate Attainment*, Journal of Public Economics 91, Issues 5-6, (June), pp. 877-899.

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

Carneiro, P. and J. Heckman (2002), *The evidence on credit constraints in post-secondary schooling*, Economic Journal, Vol. 112, pp. 705-734.

Carneiro, P. and J. Heckman (2003), *Human capital policy* in J. Heckman, A. Krueger and B. Friedman (eds.), *Inequality in America: What role for human capital policies?*, MIT Press, Cambridge, MA

Chowdry, H., Crawford, C., Dearden, L., Goodman, A., and A. Vignoles (2008). Understanding the determinants of participation in higher education and the quality of institute attended: analysis using administrative data. Institute of Fiscal Studies, working paper.

CMPO (2006), *Family background and child development up to age 7 in the Avon Longitudinal Survey of Parnets and Children (ALSPAC)*, DfES Research Report No. RR808A

Cunha, F. and J. Heckman (2007), *The Technology of Skill Formation*, American Economic Review, Vol. 92, No. 2, pp.31-47

Cunha, F., J. Heckman, L. Lochner and D. Masterov (2006), *Interpreting the evidence on life cycle skill formation* in E. Hanushek and F. Welch (eds.), *Handbook of the Economics of Education*, Chapter 12, pp.697-812, North-Holland, Amsterdam

Dearing, R. (1997). Higher education in the learning society: National Committee of Inquiry into Higher Education. London: HMSO.

Demack, S., D. Drew and M. Grimsley (2000), *Minding the Gap: Ethnic, Gender and Social Class Differences in Attainment at 16, 1988-95*, Race Ethnicity and Education, Vol. 3, No. 2, pp. 112-141

Feinstein, L. (2003), *Inequality in the Early Cognitive Development of British Children in the 1970 Cohort*, Economica, Vol. 70, No. 277, pp. 73-98

Hartog, J., Pfann, G. and Ridder, G. (1989). (Non-)Graduation and the earning function. An inquiry on self-selection. *European Economic Review* **33**, pp. 1371–1395.

Haveman, R. and Wilson, K. (2005) Economic Inequality in College Access, Matriculation and Graduation, mimeo. Paper presented to the Maxwell School Conference on Economic Inequality and Higher Education: Access, Persistence and Success, September 23-24.

HEFCE (1999) *Performance Indicators in Higher Education*, Guide 99/66, Bristol, HEFCE.

HEFCE (2005), *Young Participation in Higher Education*, HEFCE, Bristol (see: www.hefce.ac.uk/pubs/hefce/2005/05_03/05_03.pdf)

Johnes, G., and McNabb, R. (2004). Never give up on the good times: student attrition in the UK. *Oxford Bulletin of Economics and Statistics*, 66, 23-47.

Johnes, J. and Taylor, J. (1989) Undergraduate non-completion rates: differences between UK universities. *Higher Education*, 18, pp.209-225.

Machin, S. and A. Vignoles (2004), *Educational Inequality: The Widening Socio-Economic Gap*, Fiscal Studies, Vol. 25, pp. 107-28

Manski, C.F. (1989). Schooling as experimentation: a reappraisal of the postsecondary dropout phenomenon. *Economics of Education Review* 8, pp. 305–312.

McGivney, V. (1996). Staying or Leaving the Course: Non-Completion and Retention of Mature Students in Further and Higher Education. National Institute of Adult Continuing Education, Leicester, UK.

National Audit Office (2007) Staying the course: The retention of students in higher education, London: The Stationary Office.

Palmer, J. (2001) "Student Drop-out: a case study in new managerialist policy", *Journal of Further and Higher Education*, 25(3), 349-357.

Quinn, J. (2004) Understanding Working-Class 'Drop-out' from Higher Education through a Socio-cultural Lens: Cultural narratives and local contexts, International Studies in Sociology of Education, 14(1), 57-74.

Smith, J.P., & Naylor, R.A. (2001). Dropping out of university: a statistical analysis of the probability of withdrawal for UK university students. *Journal of Royal Statistical Society: Series A*, 164, 389-405.

Tinto, V., (1975). Dropouts from higher education: a theoretical synthesis of recent research. *Review of Education Research* 45, pp. 89–125.

Tinto, V. (1993) Leaving college: rethinking the causes and cures of student attrition (2nd eds) (Chicago, IL, University of Chicago Press).

Turner, S. (2008) Common Across the Atlantic: The Under-representation of Low-Income Students in Higher Education in the U.S. and the U.K., mimeo produced for The Sutton Trust Conference on Social Mobility, 1 June 2008.

Yorke, M. (1998) Non-Completion of Undergraduate Study: Some Implications for Policy in Higher Education, *Journal of Higher Education Policy and Management*, v20 n2 p189-201.

Appendix
The socio-economic gradient in HE non-continuation by group

Variables	Female	Male	Russell	Blacks
	dropout	dropout	dropout	dropout
4th deprivation quintile	-0.0069	-0.0053	-0.0035	-0.0037
	[0.0027]**	[0.0029]	[0.0027]	[0.0062]
3rd deprivation quintile	-0.0066	-0.0070	-0.0048	-0.0145
	[0.0029]*	[0.0029]*	[0.0025]	[0.0051]**
2nd deprivation quintile	-0.0112	-0.0079	-0.0065	0.0181
	[0.0030]**	[0.0031]*	[0.0031]*	[0.0129]
Lowest deprivation quintile	-0.0129	-0.0104	-0.0086	0.0120
	[0.0031]**	[0.0033]**	[0.0030]**	[0.0160]
2nd OA education quintile	0.0023	-0.0018	-0.0040	0.0013
	[0.0032]	[0.0035]	[0.0019]*	[0.0146]
3rd OA education quintile	-0.0013	-0.0060	-0.0022	-0.0123
	[0.0025]	[0.0032]	[0.0024]	[0.0092]
4th OA education quintile	-0.0038	-0.0113	-0.0016	-0.0033
	[0.0028]	[0.0029]**	[0.0027]	[0.0111]
Highest OA education quintile	-0.0066	-0.0103	-0.0053	-0.0105
	[0.0030]*	[0.0029]**	[0.0027]*	[0.0104]
Other White	-0.0080	-0.0085	0.0010	
	[0.0047]	[0.0044]	[0.0051]	
Black African	-0.0254	-0.0276	-0.0111	-0.0067
	[0.0036]**	[0.0032]**	[0.0032]**	[0.0056]
Black Caribbean	-0.0222	-0.0193	-0.0049	
	[0.0048]**	[0.0040]**	[0.0055]	
Other Black	-0.0261	-0.0261	-0.0042	-0.0105
	[0.0048]**	[0.0067]**	[0.0117]	[0.0065]
Indian	-0.0299	-0.0263	-0.0079	
	[0.0020]**	[0.0023]**	[0.0031]**	
Pakistani	-0.0258	-0.0165	-0.0025	
	[0.0030]**	[0.0033]**	[0.0041]	
Bangladeshi	-0.0198	-0.0233	-0.0078	
	[0.0045]**	[0.0044]**	[0.0057]	
Chinese	-0.0322	-0.0343		
	[0.0032]**	[0.0027]**		
Other Asian	-0.0196	-0.0142	0.0059	
	[0.0060]**	[0.0076]	[0.0076]	
Mixed ethnicity	-0.0198	-0.0174	0.0017	
	[0.0051]**	[0.0061]**	[0.0060]	
Other ethnicity	-0.0246	-0.0146	-0.0144	
	[0.0029]**	[0.0050]**	[0.0015]**	
Unknown ethnicity	-0.0046	0.0030	-0.0013	
	[0.0032]	[0.0041]	[0.0023]	
Date of birth	0.0000	0.0000	0.0000	0.0000

	[0.0000]	[0.0000]**	[0.0000]*	[0.0000]
Male			0.0005	-0.0055
			[0.0017]	[0.0053]
Disabled	-0.0087	-0.0071	-0.0017	-0.0124
	[0.0027]**	[0.0037]	[0.0024]	[0.0055]*
Level 3 at 18	-0.0285	-0.0164	-0.0025	-0.0240
	[0.0057]**	[0.0043]**	[0.0071]	[0.0120]*
Russell Group university	-0.0131	-0.0058		-0.0083
	[0.0048]**	[0.0055]		[0.0087]
RAE score of 1	0.0191	0.0042		0.0686
	[0.0162]	[0.0112]		[0.0549]
1<=RAE<2	0.0529	0.0285		0.0192
	[0.0185]**	[0.0120]*		[0.0168]
2<=RAE<3	-0.0042	-0.0032		-0.0195
	[0.0056]	[0.0051]		[0.0080]*
3<=RAE<4	-0.0011	0.0013		-0.0024
	[0.0047]	[0.0051]		[0.0106]
RAE>=5	0.0407	0.0208	-0.0006	0.0640
	[0.0199]*	[0.0189]	[0.0032]	[0.0250]*
Subject: Allied to medicine	0.0102	-0.0096	0.0020	-0.0043
	[0.0113]	[0.0104]	[0.0054]	[0.0211]
Subject: Biological science	0.0106	-0.0124	-0.0028	-0.0008
	[0.0109]	[0.0098]	[0.0050]	[0.0217]
Subject: Veterinary science	0.0141	-0.0088	0.0169	
	[0.0179]	[0.0219]	[0.0222]	
Subject: Agriculture	0.0313	0.0308	-0.0048	0.1300
	[0.0181]	[0.0220]	[0.0083]	[0.0900]
Subject: Physics	-0.0012	-0.0089	-0.0016	0.0138
	[0.0094]	[0.0096]	[0.0036]	[0.0396]
Subject: Mathematics	0.0125	-0.0148	0.0013	0.0303
	[0.0125]	[0.0078]	[0.0038]	[0.0680]
Subject: Computer science	0.0287	-0.0079	0.0119	0.0133
	[0.0165]	[0.0107]	[0.0095]	[0.0303]
Subject: Engineer & Technology	0.0254	-0.0055	0.0032	0.0137
	[0.0149]	[0.0117]	[0.0067]	[0.0322]
Subject: Architecture	0.0231	-0.0114	-0.0063	
	[0.0174]	[0.0100]	[0.0050]	
Subject: Social studies	0.0141	-0.0046	-0.0020	-0.0092
	[0.0112]	[0.0114]	[0.0039]	[0.0174]
Subject: Law	0.0073	-0.0089	0.0006	0.0200
	[0.0104]	[0.0106]	[0.0052]	[0.0279]
Subject: Business	0.0248	-0.0051	-0.0041	0.0174
	[0.0125]*	[0.0113]	[0.0037]	[0.0274]
Subject: Mass communication	0.0256	0.0018	-0.0026	0.0503
	[0.0155]	[0.0137]	[0.0043]	[0.0421]
Subject: Languages	0.0129	-0.0043	0.0023	0.0234
	[0.0109]	[0.0108]	[0.0047]	[0.0355]
Subject: History	0.0086	-0.0111	-0.0009	0.0207

	[0.0127]	[0.0112]	[0.0063]	[0.0532]
Subject: Arts	0.0484	0.0073	0.0081	0.0496
	[0.0170]**	[0.0145]	[0.0103]	[0.0415]
Subject: Education	0.0100	-0.0019	0.0170	0.0032
	[0.0112]	[0.0144]	[0.0205]	[0.0292]
Subject: Combined	0.0214	0.0030	0.0019	0.0044
	[0.0158]	[0.0134]	[0.0075]	[0.0252]
2nd KS2 quintile	-0.0006	0.0005	-0.0044	0.0091
	[0.0039]	[0.0051]	[0.0040]	[0.0109]
3rd KS2 quintile	-0.0004	0.0018	-0.0086	0.0063
	[0.0043]	[0.0051]	[0.0030]**	[0.0123]
4th KS2 quintile	0.0034	0.0087	-0.0076	0.0244
	[0.0049]	[0.0060]	[0.0041]	[0.0144]
Highest KS2 quintile	-0.0008	0.0109	-0.0103	0.0067
	[0.0045]	[0.0056]	[0.0059]	[0.0140]
2nd KS3 quintile	0.0047	-0.0105	-0.0106	-0.0023
	[0.0050]	[0.0043]*	[0.0034]**	[0.0078]
3rd KS3 quintile	-0.0009	-0.0109	-0.0103	-0.0156
	[0.0049]	[0.0045]*	[0.0040]*	[0.0096]
4th KS3 quintile	0.0054	-0.0121	-0.0078	-0.0060
	[0.0055]	[0.0049]*	[0.0062]	[0.0121]
Highest KS5 quintile	0.0067	-0.0104	-0.0148	-0.0086
	[0.0055]	[0.0061]	[0.0127]	[0.0143]
2nd capped KS4 quintile	-0.0132	-0.0051	0.0007	-0.0157
	[0.0057]*	[0.0070]	[0.0107]	[0.0095]
3rd capped KS4 quintile	-0.0251	-0.0255	-0.0048	-0.0307
	[0.0052]**	[0.0044]**	[0.0071]	[0.0129]*
4th capped KS4 quintile	-0.0350	-0.0363	-0.0104	-0.0340
	[0.0063]**	[0.0053]**	[0.0058]	[0.0152]*
Highest capped KS4 quintile	-0.0572	-0.0537	-0.0254	-0.0404
	[0.0107]**	[0.0079]**	[0.0221]	[0.0139]**
KS5 points	-0.0002	-0.0002	-0.0001	-0.0001
	[0.0000]**	[0.0000]**	[0.0000]**	[0.0000]**
Observations	72839	55584	31508	3237
Log-likelihood	-13742.1428	-10269.6963	-2805.4813	-539.704
Pseudo R-squared	0.1531	0.1633	0.0812	0.2507

Note: No institutional fixed effects.

Interactions between university status and socio-economic deprivation

Variables	All
4th deprivation quintile	-0.0069
	[0.0021]**
3rd deprivation quintile	-0.0057
	[0.0021]**
2nd deprivation quintile	-0.0081
	[0.0024]**
Lowest deprivation quintile	-0.0109
	[0.0024]**
2nd OA education quintile	0.0002
	[0.0026]
3rd OA education quintile	-0.0047
	[0.0020]*
4th OA education quintile	-0.008
	[0.0022]**
Highest OA education quintile	-0.0084
	[0.0024]**
Good university	0.0234
	[0.0047]**
Interaction with good university index	
4th deprivation quintile x good university	0.0041
	[0.0052]
3rd deprivation quintile x good university	-0.0044
	[0.0042]
2nd deprivation quintile x good university	-0.0058
	[0.0047]
Lowest deprivation quintile x good university	-0.0031
	[0.0052]
2nd OA education quintile x good university	0.0011
	[0.0046]
3rd OA education quintile x good university	0.0084
	[0.0060]
4th OA education quintile x good university	0.008
	[0.0059]
Highest OA education quintile x good university	0.0024
	[0.0053]
Institution fixed effects	Yes
Observations	127916
Log-likelihood	-

	23600.786
Pseudo R-squared	0.1711

Interactions between university status and ethnicity

Variables	All
Other White	-0.0121
	[0.0037]**
Black African	-0.0283
	[0.0020]**
Black Caribbean	-0.0222
	[0.0031]**
Other Black	-0.0279
	[0.0033]**
Indian	-0.0277
	[0.0013]**
Pakistani	-0.0221
	[0.0021]**
Bangladeshi	-0.0193
	[0.0031]**
Chinese	-0.0328
	[0.0022]**
Other Asian	-0.0244
	[0.0040]**
Mixed ethnicity	-0.0193
	[0.0035]**
Other ethnicity	-0.0196
	[0.0029]**
Unknown ethnicity	0.0003
	[0.0030]
Good university	0.0185
	[0.0040]**
Interaction with good university index	
Other White x good university	0.0200
	[0.0119]
Black African x good university	0.0313
	[0.0177]
Black Caribbean x good university	0.0104
	[0.0135]
Other Black x good university	0.0463
	[0.0414]
Indian x good university	0.0202
	[0.0120]
Pakistani x good university	0.0300
	[0.0135]*
Bangladeshi x good university	0.0003
	[0.0207]
Chinese x good university	0.0136
	[0.0256]
Other Asian x good university	0.0676
	[0.0320]*

Mixed ethnicity x good university	0.0096
	[0.0215]
Other ethnicity x good university	-0.0041
	[0.0094]
Unknown ethnicity x good university	-0.0045
	[0.0032]
Institution fixed effects	Yes
Observations	127916
Log-likelihood	-23590.114
Pseudo R-squared	0.1714