Impact and Effectiveness of Widening Access to HE in Wales
Working Paper Series[1] WAQNCW2014-3

# Widening Access to higher education in Wales: Analysis using linked administrative data. Caroline Wright

[1] **Working Papers are intended to present initial analysis and findings. Please do not cite them without permission of the author(s).**

# Foreword

The Wales Institute of Social and Economic Research, Data and Methods (WISERD) is currently evaluating the impact and effectiveness of Widening Access to higher education (HE) in Wales. This project is funded through the Economic and Social Research Council's (ESRC) Secondary Data Analysis Initiative (ES/K004247/1); and by the Higher Education Funding Council for Wales (HEFCW). It will be completed by December 2014.

Access to higher education has become an extremely controversial area of policy, as successive UK administrations have sought to balance increasing student fees with ensuring that HE is open to individuals from as wide a range of social backgrounds as possible. Moreover, relatively distinctive approaches have been adopted in the different devolved administrations of the UK. For example, currently, the Welsh Government has undertaken to pay the increased costs to students arising from the abolition of the fees cap. However, the evidence-base for evaluating different approaches to widening access is relatively weak. Accordingly, WISERD, the HEFCW and the Welsh Government (WG) are collaborating on this innovative research study.

The research analyses how individuals who are resident in Wales progress through secondary school, into sixth forms and further education colleges for post-16 education and on to HE. It also explores what are the key factors here in determining whether individuals progress through the education system to HE or not. What are the relative impacts of the social characteristics of individuals, their previous educational attainment and their progression through the education system? What does this imply for the effects of barriers at the point of entry to HE, such as fees levels, entry processes and so forth? Answers to these questions are known for England, but not for other parts of the UK.

The analysis is based on the innovative use of three linked sources of information, the data for each of which are collected initially for administrative purposes. These are: the National Pupil Database (NPD) for Wales; the Lifelong Learning Wales Record (LLWR); and Higher Education Statistics Agency (HESA) data. By linking these together, it is possible to trace individual trajectories through the education system to entry to HE. It is also possible to compare systematically the trajectories of those who do participate in HE with those who do not. Moreover, using sophisticated statistical techniques, it is possible to determine which are the most influential factors in shaping patterns of HE participation. Results here will be compared with those that have been produced by similar analyses in England.

A second part of the proposed study (funded by additional resources made available by the HEFCW) investigates the development of distinctive approaches to widening access to HE by successive Welsh administrations since devolution in 1999. Of key significance here is to establish the rationales that underpin the approaches adopted in Wales; and to compare these with those that have informed policy approaches in the other countries of the UK and England, in particular. In addition, the study examines the ways in which national policies have been implemented by the Welsh universities, paying special attention to the assumptions about the determinants of HE participation that are in play here. This part of the study is based on fieldwork, comprising the analysis of official and semi-official documents and interviews with politicians and senior officials responsible for widening access policies; and with the professionals inside the universities responsible for implementing these policies.

The results of the research will be fed directly into the deliberations of the WG and the HEFCW on the future development of policies on widening access to HE, which will be especially intensive over the next few years. Moreover, they will also provide the basis for working with the professionals in the universities with responsibility for implementing widening access policies, to integrate the use of analyses of administrative data more firmly into their day-to-day practices.

# Widening Access to higher education in Wales: Analysis using linked administrative data

## Introduction

This paper aims to replicate analysis conducted by Chowdry et al. (2013) which considers the determinants of participation in higher education (HE) in England. Using linked individual-level data, for both participants and non-participants in HE, Chowdry et al. track 2 cohorts of young people from age 11 through to age 20. They demonstrate that pupils from lower socio-economic status (SES) backgrounds are much less likely to participate in HE, than pupils from higher SES backgrounds. However, they also show that this difference does not emerge at the point of entry to HE, rather it is that lower SES pupils do not achieve as highly as higher SES pupils throughout education and subsequently do not have the same options available to them. They also show that socio-economic difference in participation does remain on entry to HE even allowing for prior attainment, although this is much smaller.

The first part of this paper will follow the methodology of the Chowdry paper closely given the data available to this project and in a similar way will assess the relative contributions of SES background and prior academic attainment on participation in HE, for the Welsh population. Whilst comparisons between Chowdry et al.'s work (England) and this work (Wales) will be made, it is recognised that this should be done cautiously, given the subtle differences in the data used.

The second part of the paper will extend the previous analysis by adopting an exploratory, multilevel approach. The reasons for using a multilevel approach will be explored in more detail in the Methods section, but to summarise the approach taken in this extension will allow for the exploration of other potentially important contexts e.g. local authorities. This work will both extend the current research context (this type of work has not previously been conducted using Welsh data), but will also allow us to additionally explore what makes schools and/or local authorities (un)successful in this domain a finding which could be potentially very important for policy formulation.

## Data

The data used in this paper is based on three linked administrative data-sets: the National Pupil Database (NPD) for Wales including Pupil Level Annual Schools Census (PLASC) data; individual learner records from the Lifelong Learning Wales Record (LLWR) for young people who are registered at post-compulsory educational institutions (not including school sixth forms); and individual student records from the Higher Education Statistics Agency (HESA). At its core, the database follows the educational trajectories of 3 cohorts of young people who were in Year 11 (the final year of compulsory schooling, referred to as Key Stage 4) during 2004/5, 2005/6 and 2006/7. The Welsh Government procured the matching of attainment data from school (NPD) and further education (LLWR) individual learner records for Wales from 2002-3 to 2008-9. In addition, these data have been matched to HE participation data (HESA) for 2007-8 and 2008-9, each with all three year-cohorts of undergraduate students included. Thus, data are available for four years of first-year entrants to HE from 2006-7 to 2009-10. Look-up tables of individual identifiers are available, permitting further linking and analysis of the data-sets. These data sources provide the basis for the analysis of individual trajectories through the compulsory education sector (from the age of 11), to post-16 education and on to HE.

## Data coverage

The data provide a census of state school children and young people in Wales and include academic outcomes in the form of public examination results at age 16 (GCSEs). Data is available regarding A-Level scores, but only for the HE participating population and as such can only be used in a limited way. The data also include a variety of additional individual characteristics including date of birth, home postcode, gender, ethnicity and entitlement to free school meals (FSM) a proxy for socio-economic disadvantage. Unlike the Chowdry paper we do not have public examinations results at age 16 and 18 for those outside the state school sector, although we do have A-level results/tariff score for those who made it to HE.

This data also contains information regarding whether or not each student enrolled in HE by age 19 or 20. Like Chowdry *et al*., we do not know whether an individual applied to participate in HE through UCAS, rather we only know whether they were admitted. This means that the analysis is limited in the sense that it can’t determine whether the HE participation rate of particular group(s) is low because they don't apply to HE or if they are not successful during the application/admissions process.

### Outcome Variable

The analysis in this paper will focus on participation in HE, as defined by those who enrol in an HE program by age 20.

### Measuring socio-economic background

In this analysis we have relied upon information regarding students' eligibility for FSM at age 15 as a proxy measure of students' socio-economic background. Although this measure is prevalent in educational research, it can be problematic given its potential to exclude large numbers of children and young people who don't claim FSMs but nonetheless experience high levels of poverty. Hobbs and Vignoles (2010) found that, although the majority of pupils claiming FSM do live in low-income families, not all pupils living in low-income households claim FSM. Thus, FSM is only a partial measure of deprivation and as such the actual effect of socio-economic disadvantage may be much larger than estimated in this research. The use of NS-SEC social classifications, in addition to FSM data, as explanatory variables would therefore have been preferable, but these are only available for individuals who eventually progress to HE and even amongst this population there is still missing data to consider. We have the home postcode of all the young people in the data set, which can be used to link in detailed information regarding the neighbourhood in which they live. However, once again, this isn't without its problems. For example, Taylor et al. (2013) demonstrate that coming from a low participation neighbourhood (LPN) does not necessarily mean that students are from socio-economically disadvantaged households, nor that they are necessarily non-traditional or first-generation entrants to HE (p.158).

Indeed, it would have been preferable to consider a more nuanced form of socio-economic background. For example, a recent paper by Bukodi and Goldthorpe (2013) finds that socio-economic status (SES), family income and parental education all have significant separate effects, even when modelled together. This is important because these effects are often used interchangeably or as proxies for one another, but in fact it is shown that they have 'independent and distinctive effects on educational attainment' (p.1024).

We will use the Welsh Index of Multiple Deprivation (WIMD) to provide a broader indicator of deprivation. The WIMD is a relative measure of concentrations of deprivation at the lower layer super output area (LSOA) level. In this respect, deprivation is understood as a wider concept than poverty, encompassing measures of both economic and social resources and opportunities. As such, WIMD is constructed from eight different types of deprivation. These are: income; housing; employment; access to services; education; health; community safety; and physical environment (http://wales.gov.uk/statistics-and-research/welsh-index-multiple-deprivation/?lang=en). Again, this approach is not without its' problems: using aggregate data falls foul of the ecological fallacy, which is the fallacious attribution of aggregate results on to individuals (Robinson, 1950), thus it is important when describing this population to distinguish these young people as 'those living in deprived areas' rather than 'deprived young people'.

Whilst Chowdry *et al*. used principal components analysis (PCA) to combine various measures of a young person's socio-economic circumstances (FSM status; index of multiple deprivation (IMD) score; a classification of residential neighbourhood type; and three very local area-based measures: specifically the proportion of individuals in each area who work in managerial or professional occupations, whose highest educational qualification is level 3 or above and who own their home). At this stage there is no plan to replicate the PCA. It was decided that using the WIMD measure along with an indicator of FSM would be sufficient for our purposes.

Table 1 details the HE participation rates, according to the Welsh Index of Multiple Deprivation (WIMD) and gender. The first row is for the overall WIMD measure, this is followed by each of the constituent parts of the WIMD (employment, income, education, health, access to services, physical environment, housing and community safety). It shows that females are more likely to participate in HE than males, regardless of their WIMD classification, this is true of the combined WIMD as well as all its constituent parts. However, the gender gap tends to be greatest in the first and second quintiles, respectively. While all the other WIMD classifications indicate a negative relationship associated with living in low WIMD neighbourhoods and HE participation, the measure related to 'access to services' indicates the converse – those living in LSOAs with the lowest accessibility scores, possibly rural communities, have the highest participation rates and those living in high accessibility areas have the lowest participation rates. The sixth column in Table 1 shows the participation gap between the top and bottom WIMD quintiles. Interestingly, the gap is largest for females in the top and bottom quintiles, and this is the case for every one of the WIMD classifications. So whilst females from the bottom quintile are still more likely to participate than males from the same quintile, there is a greater differential between them and their more privileged female contemporaries than between equivalent male students. It is also noteworthy that the overall WIMD measure, WIMD education and WIMD income precipitated the largest gaps in participation, whereas the WIMD measure relating to physical environment made almost no difference to young peoples' participation.

### Other individual characteristics

In addition to WIMD quintile, our models will include a number of other individual-level characteristics in line with the Chowdry paper. These include season of birth, ethnicity, special educational need indicator (SEN), free school meal take up (FSM) and cohort. We do not use English as an additional language as this was not available in the NPD data. We also include test scores from age 16. The intention had been to include test scores from earlier stages of education, however the data for these was incomplete and unusable.

Table 2 details HE participation rates, according to gender and a number of other individual characteristics. This shows that females have a higher overall participation rate (38.7%) than males (30.3%). It also shows that non-White students have the highest participation rates, followed by White Other, White British and then 'don't know/not sure', this is the case for both males and females. There are marked differences in participation rates between those who claim FSM and those who don't, with a differential of 21.3% for males and 27.1% for females. Unsurprisingly, those with no SEN have the highest participation rates – 35.4% (males) and 42% (females), and those with a statement of SEN have the lowest – 7.9% (males and 7.4% (females). The cohort for 2006 has the highest participation rate and 2005 the lowest for both gender groups. Finally, there doesn't appear to be any discernible difference in participation rate according to season of birth, this is the case for both males and females.

### Table 1: HE participation rates by WIMD quintile and gender

| | Top (1st) quintile (%) | 2nd quintile (%) | Middle quintile (%) | 4th quintile (%) | Bottom (5th) quintile (%) | Difference (1st - 5th) (percentage points) | Overall (%) |
|---|---|---|---|---|---|---|---|
| **WIMD** | | | | | | | |
| **Male** | 48.8 | 35.7 | 31.4 | 23.9 | 16.3 | 32.5 | 30.5 |
| **Female** | 57.9 | 46.8 | 40.9 | 30.6 | 21.9 | 36.0 | 38.9 |
| **WIMD (Employment)** | | | | | | | |
| **Male** | 45.6 | 36.4 | 29.7 | 23.2 | 17.3 | 28.3 | 30.5 |
| **Female** | 55.9 | 47.4 | 37.9 | 30.1 | 23.1 | 32.8 | 38.9 |
| **WIMD (Income)** | | | | | | | |
| **Male** | 48.4 | 36.8 | 30.9 | 23.7 | 16.3 | 32.2 | 30.5 |
| **Female** | 57.9 | 48.3 | 39.7 | 31.3 | 21.2 | 36.7 | 38.9 |
| **WIMD (Education)** | | | | | | | |
| **Male** | 50.2 | 36.5 | 30.6 | 23.5 | 15.7 | 34.5 | 30.5 |
| **Female** | 59.6 | 48.1 | 39.8 | 30.5 | 20.8 | 38.8 | 38.9 |
| **WIMD (Health)** | | | | | | | |
| **Male** | 44.8 | 36.7 | 30.4 | 24.2 | 18.8 | 26.0 | 30.5 |
| **Female** | 54.8 | 46.4 | 39.9 | 31.2 | 24.8 | 30.1 | 38.9 |
| **WIMD (Access to services)** | | | | | | | |
| **Male** | 26.9 | 26.4 | 27.1 | 34.3 | 38.0 | -11.1 | 30.5 |
| **Female** | 33.9 | 33.4 | 34.4 | 43.3 | 49.7 | -15.8 | 38.9 |
| **WIMD (Physical environment)** | | | | | | | |
| **Male** | 31.5 | 32.0 | 31.3 | 30.7 | 26.9 | 4.6 | 30.5 |
| **Female** | 40.1 | 41.4 | 39.1 | 39.3 | 34.3 | 5.9 | 38.9 |
| **WIMD (Housing)** | | | | | | | |
| **Male** | 44.7 | 30.6 | 27.9 | 27.1 | 22.3 | 22.4 | 30.5 |
| **Female** | 53.9 | 39.5 | 36.8 | 34.5 | 29.4 | 24.5 | 38.9 |
| **WIMD (Community safety)** | | | | | | | |
| **Male** | 44.1 | 38.7 | 28.2 | 24.5 | 18.7 | 25.4 | 30.5 |
| **Female** | 55.5 | 48.1 | 38.2 | 30.9 | 23.5 | 31.9 | 38.9 |

### Table 2: Participation rates by personal characteristics

| Participation in HE by age 20 | | | | |
|---|---|---|---|---|
| | Male | | Female | |
| | No. | % | No. | % |
| **Overall** | **16,978** | **30.3** | **21,082** | **38.7** |
| **Ethnicity** | | | | |
| **White British** | 15,592 | 30.0 | 19,442 | 38.3 |
| **White Other** | 298 | 35.6 | 369 | 45.7 |
| **Non-White** | 677 | 40.3 | 791 | 48.0 |
| **DK/NS** | 411 | 28.4 | 480 | 37.1 |
| **FSM Status** | | | | |
| **Yes** | 930 | 12.0 | 1198 | 15.4 |
| **No** | 16,048 | 33.3 | 19,884 | 42.5 |
| **SEN status** | | | | |
| **Nothing** | 15,834 | 35.4 | 20,371 | 42.0 |
| **Action** | 678 | 12.1 | 448 | 13.3 |
| **Action Plus** | 247 | 8.7 | 180 | 11.9 |
| **Statement** | 219 | 7.9 | 83 | 7.4 |
| **Cohort** | | | | |
| **2005** | 5436 | 28.6 | 6562 | 35.8 |
| **2006** | 5771 | 31.4 | 7270 | 40.5 |
| **2007** | 5771 | 31.1 | 7250 | 39.8 |
| **SOB** | | | | |
| **Autumn** | 4324 | 30.8 | 5281 | 39.2 |
| **Winter** | 4065 | 30.3 | 5162 | 38.9 |
| **Spring** | 4353 | 30.6 | 5272 | 38.5 |
| **Summer** | 4236 | 29.6 | 5367 | 38.0 |

Table 3 details HE participation rates, according to gender and two attainment measures: GCSE CSI and GCSE points band. It is recognised that as well as level of prior attainment having an impact, the subjects studied may too have an effect, for this reason the Core Subject Indicator (CSI) for GCSE results, which identifies those students who achieved 5 GCSEs at grades A*-C including Maths and English, is also included. The results in Table 3 show that those who have achieved GCSE CSI have a higher overall participation rate than those who don't the difference between the two groups is 52.7% (males) and 55.5% (females). It also shows that the higher your GCSE points score is, the more likely you are to participate in HE.

### Table 3: Participation rates by attainment characteristics

| Participation in HE by age 20 | | | | |
|---|---|---|---|---|
| | Male | | Female | |
| | No. | % | No. | % |
| **GCSE CSI** | | | | |
| **Yes** | 13,084 | 63.7 | 16,342 | 70.6 |
| **No** | 3894 | 11.0 | 4740 | 15.1 |
| **GCSE Points Band** | | | | |
| **0-50** | 4603 | 12.1 | 4027 | 13.2 |
| **50-55** | 2351 | 48.8 | 2680 | 48.8 |
| **55-60** | 2475 | 62.6 | 3245 | 64.0 |
| **60-65** | 2320 | 73.1 | 3130 | 73.3 |
| **65-70** | 1936 | 82.0 | 2773 | 81.2 |
| **70-75** | 1368 | 84.7 | 2226 | 88.4 |
| **75+** | 1925 | 90.5 | 3001 | 90.8 |

## Methods

For the purposes of replication, we are following the methods adopted by Chowdry et al. (2013)[2]. They are clear that they wish to incorporate a 'school effect' and discuss the advantages and disadvantages of adopting fixed and random effects models to achieve this. Whilst they use linear regression with fixed effects models to estimate their final models and derive their results, they additionally estimate logistic models for participation in HE, with school random effects as a 'check for robustness'. Thus our research will follow this approach and produce both types of models. It is important to note that, using a fixed effects approach is equivalent to adding a dummy for each school. This may be problematic when your sample has small clusters, i.e. there are not many students within each school. However, this is justified by Chowdry et al., on the basis that the data used is census data, where students are nested in schools and schools within two cohorts of data. Similarly, our data reflects students nested in schools, nested in three cohorts of data.

Ordinarily, with a binary outcome variable, logistic regression is adopted, however, because of the large number of fixed effects that need to be included in the model, this is not possible and a linear regression is used. Chowdry et al. argue that,

> there are circumstances under which the linear probability model provides a close approximation of the logit model, namely where the probability of participation is between 0.25 and 0.75 (p.443).

[2] For a more detailed discussion regarding why this particular approach was chosen, please see the original paper.

In line with the Chowdry *et al*. (2013) paper, our models will be estimated sequentially, separately for females and males, and will begin with a model only including our indicator variable of SES background (WIMD quintile) and a cohort dummy variable. This provides an estimate of the underlying differences in HE participation by SES. The following model includes a number of individual covariates (ethnicity, FSM a special educational needs indicator and season of birth), as well as school fixed effects. Following this we are able to estimate the extent to which these SES differences may be explained by differences in prior academic attainment through the inclusion of GCSE. This method will allow us to estimate both the extent to which SES affects HE participation directly, as well as indirectly, through its effect on prior academic attainment.

The results in Table 4 (males) and

Table 5 (females) show that significant differences in HE participation exist between students who live in different WIMD quintiles, before any covariates have been included in the model. These differences diminish as the individual covariates and the school effects are included, as the omitted variable bias reduces. However, these differences between WIMD quintiles do not disappear entirely and in fact remain highly significant. This reflects the results found in the Chowdry paper.

### Table 4: Gradients in HE participation for males: results from a logistic model with random school effects

| Statistic | No covariates | Individual covariates and school REs | Plus GCSE points score | Plus GCSE CSI |
|---|---|---|---|---|
| **2nd WIMD quintile** | -0.539**(0.027) | -0.448**(0.029) | -0.229**(0.037) | -0.318**(0.036) |
| **Middle WIMD quintile** | -0.730**(0.027) | -0.591**(0.030) | -0.252**(0.038) | -0.402**(0.036) |
| **4th WIMD quintile** | -1.111**(0.029) | -0.974**(0.030) | -0.358**(0.039) | -0.648**(0.036) |
| **Bottom WIMD quintile** | -1.583**(0.030) | -1.316**(0.033) | -0.448**(0.043) | -0.849**(0.038) |
| **Observations** | **55465** | **55465** | **55465** | **55465** |
| **Clusters** | | **127** | **127** | **127** |
| **P-value** | | 0.000 | 0.000 | 0.000 |

**Table 5: Gradients in HE participation for females: results from a logistic model with random school effects**

| Statistic | No covariates | Individual covariates and school REs | Plus GCSE points score | Plus GCSE CSI |
|---|---|---|---|---|
| **2nd WIMD quintile** | -0.445**(0.028) | -0.375**(0.030) | -0.237**(0.037) | -0.282**(0.035) |
| **Middle WIMD quintile** | -0.685**(0.028) | -0.559**(0.030) | -0.256**(0.036) | -0.369**(0.036) |
| **4th WIMD quintile** | -1.137**(0.029) | -1.002**(0.031) | -0.436**(0.038) | -0.683**(0.037) |
| **Bottom WIMD quintile** | -1.589**(0.031) | -1.5329**(0.032) | -0.419**(0.040) | -0.853**(0.039) |
| **Observations** | **53994** | **53994** | **53994** | **53994** |
| **Clusters** | | **128** | **128** | **128** |
| **P-value** | | 0.000 | 0.000 | 0.000 |

**All specifications include a cohort dummy variable. The individual covariates included from the third column onwards are ethnicity, whether they have special educational needs, season of birth and whether they claim free school meals. Random school effects are also included from the third column onwards. Standard errors are clustered at the school level.**

*** Significance at the 5% level**
****Significance at the 1% level**

## Results I: Replicating Chowdry

The following section presents the results from the final models. Table 6 and Table 7 present the estimates of the impact of SES (as measured by WIMD) on HE participation in Wales by age 20, for males and females separately. The log likelihood is a measure of 'badness of fit' and as such, any reduction in this number may be understood as improvement in model fit and in the absence of an r squared value, gives us an idea of whether the model fit has improved.

The second column in these tables shows the 'raw' differences, i.e. differences in HE participation rate by WIMD quintile, without taking account of any other factors (apart from cohort). The third column takes account of individual covariates and school fixed effects. The remaining columns show how the determining effect of WIMD is altered by the inclusion of prior academic attainment this is present for GCSE as a continuous variable and for GCSE CSI.

The second columns show that there are large and significant differences in HE participation between different WIMD quintiles, whereby those living in LSOAs in the bottom quintile are the least likely to participate and those living in the highest quintile, are the most likely. Once individual covariates and school fixed effects have been taken into account, these differences are reduced, and among males there is no significant difference between the base category and the $2^{nd}$ WIMD quintile (this is not the case amongst females and a significant relationship remains for all quintiles when compared to the base category). However, although the effect of WIMD quintile has been reduced with the inclusion of individual covariates and school fixed effects (the estimates are smaller), they remain statistically significant. This finding suggests that differences in individual characteristics and the type of school attended by young people from different WIMD backgrounds provide some explanation for why students from lower WIMD backgrounds are less likely to go to university than those living in high WIMD neighbourhoods.

The third and fourth columns show how HE participation rates vary between students from different WIMD backgrounds, who otherwise have similar observable characteristics, attend the same schools and follow the same pattern of academic attainment at age 16. The inclusion of previous academic attainment into the model, results in substantial reductions in the sizes of the estimates for each of the WIMD quintile. Indeed in all cases the estimates reduce by at least half. The inclusion of GCSE score results in the largest reduction in the log likelihood and as such may be understood as a better model and this measure of prior academic attainment will therefore be used in preference to GCSE CSI.

These results suggest that WIMD and prior academic attainment (GCSE score) are highly correlated. It is unfortunate that we do not have adequate data for earlier stages of education, as it would have been interesting to estimate how early these educational trajectories are set in place.

Once all available measures of educational attainment have been added to the model, males and females in the bottom quintile are 6.7% and 7.5% less likely to participate in HE than equivalent young people in the top quintile, this is respectively around 20.6% and 20.9% of the raw differences observed in the second column. What is striking in the Chowdry paper is that the difference in participation rates between the fourth and fifth quintile of WIMD was not significant for males and just 0.6 for females. However, when these are compared in our data, it is not the bottom quintile, but the fourth quintile that are the least likely to participate (for both male and female students). Chowdry et al., found the largest differences between the top and second quintiles (2.4% for both males and females), which we also found, although the difference was greater in our data (4.4% - males and 4.6% - females). There is also quite a large difference for females between the middle and 4$^{th}$ quintile (3.1%). This indicates that the WIMD difference in HE participation is greatest at the top of the WIMD distribution. It is important to note that these slight differences between ours and Chowdry's analysis may be related to the slight differences between the two data sets.

### Table 6: Gradients in HE participation for males: results from a linear model with fixed effects

| Statistic | No covariates | Individual covariates and school FEs | Plus GCSE points score | Plus GCSE CSI |
|---|---|---|---|---|
| **2nd WIMD quintile** | -0.131** (0.006) | -0.105 (0.006) | -0.046** (0.005) | -0.056** (0.004) |
| **Middle WIMD quintile** | -0.173** (0.006) | -0.136** (0.006) | -0.053** (0.005) | -0.070** (0.004) |
| **4th WIMD quintile** | -0.249** (0.006) | -0.208** (0.006) | -0.071** (0.005) | -0.106** (0.005) |
| **Bottom WIMD quintile** | -0.324** (0.006) | -0.254** (0.006) | -0.067** (0.005) | -0.125** (0.005) |
| **Observations** | **55465** | **55465** | **55465** | **55465** |
| **Log likelihood** | **68053.637** | **64817.472** | **45630.810** | **49172.804** |
| **No. of schools** | | **127** | **127** | **127** |
| **P-value** | | 0.000 | 0.000 | 0.000 |
| **% with predicted values <0 or >1** | 70% | 66% | 54% | 67% |

### Table 7: Gradients in HE participation for females: results from a linear model with fixed effects

| Statistic | No covariates | Individual covariates and school FEs | Plus GCSE points score | Plus GCSE CSI |
|---|---|---|---|---|
| **2nd WIMD quintile** | -0.111** (0.007) | -0.089** (0.007) | -0.044** (0.006) | -0.050** (0.006) |
| **Middle WIMD quintile** | -0.169** (0.007) | -0.134** (0.007) | -0.053** (0.006) | -0.067** (0.006) |
| **4th WIMD quintile** | -0.273** (0.006) | -0.229** (0.007) | -0.084** (0.006) | -0.118** (0.006) |
| **Bottom WIMD quintile** | -0.359** (0.006) | -0.283** (0.007) | -0.073** (0.006) | -0.136** (0.006) |
| **Observations** | **53994** | **53994** | **53994** | **53994** |
| **Log likelihood** | **71806.704** | **68992.444** | **49500.799** | **52722.819** |
| **No. of schools** | | **128** | **128** | **128** |
| **P-value** | | 0.000 | 0.000 | 0.000 |
| **% with predicted values <0 or >1** | 62% | 59% | 50% | 60% |

**All specifications include a cohort dummy variable. The individual covariates included from the third column onwards are ethnicity, whether they have special educational needs, season of birth and whether they claim free school meals. Random school effects are also included from the third column onwards. Standard errors are not currently clustered at the school level.**

*** Significance at the 5% level / **Significance at the 1% level**

## Checks for robustness

Chowdry et al. highlighted the potential problems inherent in using linear regression with fixed school effects, in that it can produce predicted values that are less than 0 or more than 1. However, it is much more concerning in our research, because the percentage of results adhering to this rule exceeds 60% (where for Chowdry et al. it was between 7-17%).

As already mentioned, one check for robustness is to compare the estimates from the logistic model with random effects, with the final model (a linear model with fixed effects). Whilst it is not possible to directly compare the estimates from logistic and linear models (one represents the log odds and the other a probability), it is possible to perform an anti-logit calculation on the log odds, to produce estimates that are comparable with probabilities. This was done for the estimates of the final model and the sizes were highly consistent. It is also reassuring to see that the estimates for WIMD quintile show the same levels of significance for both sets of models and follow a similar pattern of reduction in their relative sizes with the inclusion of the other covariates. It is however, noteworthy that while the nature of all the relationships (apart from SEN: Action Plus) are the same in both the logistic and linear models, not all of the other coefficients replicate one another. This difference and others like it will be discussed later in this paper.

Chowdry et al. investigated some interactions between the SES quintiles and ethnicity and prior academic attainment. We did not find any significant results from the same interactions, suggesting that the main effects described in the final model sufficiently explain the effect of WIMD, ethnicity and attainment, respectively.

## Extending the analysis: a multilevel approach

As already mentioned, using fixed effects models to estimate the school effect is problematic for several reasons.

Firstly, by adding in a separate dummy variable for every school, the fixed effects approach simply ‘models out’ the effect of school. This may be an appropriate strategy if you are simply assessing the effectiveness of a national policy, such as widening access overall. However, if you are interested in exploring what it is about particular schools, that makes a difference to young people’s educational trajectories, then a random effects model is more appropriate.

Second, there is no single parameter to assess between-group differences and as such the fixed effects model provides no metric for school effects. This means that research using these methods, are unable to compare the size of the school effects, with the other effects in the model i.e. you don’t get any relative understanding of the influence of different factors.

Thirdly, it is not possible to include group-level predictors e.g. school type, because all degrees of freedom have been consumed at the group-level, meaning that you’re not able to explain the ways that schools produce different outcomes - this has to be done in a separate analysis.

Finally, the fixed effects approach may result in potentially unstable estimates, unless all cluster sizes are large. Clusters based on a small number of units will be highly volatile and therefore unreliable. It is also problematic to attribute equivalent meaning to schools with different sized populations.

The Chowdry analysis will be extended by including random effects for both schools and local authorities, allowing us to additionally: quantify the ‘school effect’ and as such, compare it with the size of the other effects; allow for the testing of group-level predictors in order to explain previously unexplained school- and local authority-level variance; and produce more robust and reliable estimates for the school- and local authority-effects.

## Examining the data

Table 8 shows the average HE participation rate at each local authority in Wales. The raw data shows that there is a great variation between local authorities in terms of their average participation rates. For example, the highest participation rate among males is found in the Vale of Glamorgan - 38.7%, whilst for females it is in Ceredigion - 50.4%. The lowest participation rates are found in Wrexham for both males and females and are 23.2% and 30.0% respectively. This justifies adopting a multilevel approach, in order to further explore this effect.

### Table 8: HE Participation rate by local authority

| HE Participation rate by LA (%) | | |
|---|---|---|
| | **Males** | **Females** |
| Blaenau Gwent | 23.7 | 30.1 |
| Bridgend | 31.1 | 40.7 |
| Caerphilly | 24.0 | 32.5 |
| Cardiff | 33.4 | 39.1 |
| Carmarthenshire | 32.7 | 44.3 |
| **Ceredigion** | 34.2 | **50.4** |
| Conwy | 30.9 | 40.6 |
| Denbighshire | 27.0 | 36.5 |
| Flintshire | 27.8 | 36.7 |
| Gwynedd | 32.3 | 43.4 |
| Isle of Anglesey | 32.3 | 44.0 |
| Merthyr Tydfil | 37.8 | 48.0 |
| Monmouthshire | 35.4 | 46.2 |
| Neath Port Talbot | 27.1 | 36.8 |
| Newport | 30.6 | 36.5 |
| Pembrokeshire | 32.4 | 41.9 |
| Powys | 34.4 | 43.9 |
| Rhondda | 28.7 | 35.2 |
| Swansea | 32.1 | 39.1 |
| Torfaen | 24.3 | 30.3 |
| **Vale of Glamorgan** | **38.7** | 43.9 |
| **Wrexham** | **23.2** | **30.0** |

# Results from the multilevel modelling, extending the analysis

## Individual-level Effects

The results displayed in Table 9 show that young people in the top SES quintile are the most likely to participate in HE, followed by each of the other 4 quintiles in order of disadvantage. For males, the largest gap in participation is between the top and second quintile (12%), for females, a large gap was also present between the top and second quintile (15%), but a similar size gap was also found between the 3$^{rd}$ and 4$^{th}$ quintiles (14%), suggesting that the differential participation rates occur at different thresholds for the different gender groups. In addition to the effect of WIMD quintile, there is a supplementary effect associated with individual SES as measured by FSM status. Males' who receive FSM experience less of a disadvantage than their female counterparts, but are still 12% less likely to participate than those who don't claim FSM. This compares to females who claim FSM who are 21% less likely to participate than those who don't. Both male and female students are equally affected by their GCSE points score – an additional 11% more likely to participate for each additional GCSE point scored. The effect of ethnicity on participation rates is marked in this analysis, amongst male students, 'non-

Whites' are almost 2 and a half times more likely to participate than their 'White British' contemporaries and amongst females, they are twice as likely. There is no significant cohort effect for males, but one is present for females, who see a significant increase in participation rate of 17% between 2005 and 2006. While still displaying a higher participation rate than 2005, 2007 does not have as a high a participation rate as 2006. Finally, season of birth is shown to be significant in this model and surprisingly, it is those born in summer who are the most likely to participate in HE and those born in autumn who are the least likely, this is the case for both male and females.

### Table 9: Results for all the covariates in the final model, as odds ratios (OR)

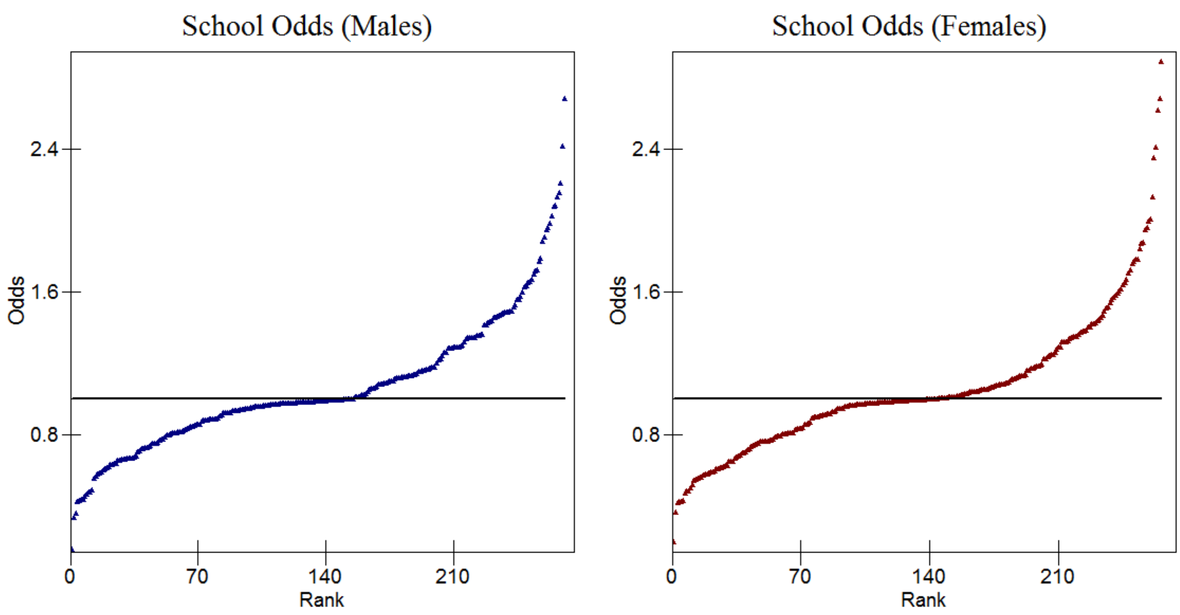
| Explanatory Variable | Final Model (males) | Final Model (females) |
|---|---|---|
| WIMD 2 | 0.88*** | 0.85*** |
| WIMD 3 | 0.85*** | 0.80*** |
| WIMD 4 | 0.75*** | 0.66*** |
| WIMD 5 | 0.66*** | 0.65*** |
| GCSE Points GM | 1.11*** | 1.11*** |
| SEN: Nothing | 0.84*** | 0.93 |
| SEN: AP | 1.02 | 1.10 |
| SEN: Statement | 1.91*** | 1.54*** |
| FSM: Yes | 0.88*** | 0.79*** |
| Ethnicity: White Other | 1.34** | 1.41*** |
| Ethnicity: Non-White | 2.38*** | 2.00*** |
| Ethnicity: DK/NS | 1.01 | 0.91 |
| 2006 | 1.03 | 1.17*** |
| 2007 | 0.99 | 1.12*** |
| Winter | 1.12*** | 1.12*** |
| Spring | 1.22*** | 1.14*** |
| Summer | 1.28*** | 1.25*** |

***Significance at the 10% level**
**** Significance at the 5% level**
*****Significance at the 1% level**

## School effects

Figure 1 shows the differences in HE participation according to school. These differences may be understood as supplementary to the already discussed individual differences. Figure 1 shows that there is substantial between-school difference in terms of young people's propensity to participate in HE. The horizontal line on both of the graphs represents the national average (an odds ratio of 1). Had there been no between-school differences, all of the residuals would fall on or close to this line. However, the residuals depart significantly from the national average and as such we can conclude that, even after attainment and a variety of other factors have been accounted for schools make a big difference to a young people's likelihood of HE participation. Additionally, we can see that for both males and females there are a small number of schools who are performing well above average. Demonstrating that young people attending the best performing schools are more than two and a half times more likely (males) and nearly three times more likely (females) to participate in HE, than an equivalent student, attending the average school. Those attending the worst performing school are 42% (males) and 43% (females) less likely to participate than those attending an average school. Interestingly, none of the highest/lowest performing school for males were the same as those for females. Figure 1 represents the *unexplained* variance that is linked to schools and as such does not give any indication about what it is about particular schools that makes them perform well or badly in terms of HE participation. This could be further explored by identifying the schools in question and undertaking additional qualitative research within the school(s) to try and identify what common factors affect a school's performance in this respect. Alternatively, we could assess (by running additional models) whether the inclusion of particular school-level variables result in a reduction in this unexplained school-level variance.

### Figure 1: Odds of participating in HE according to school for males and females

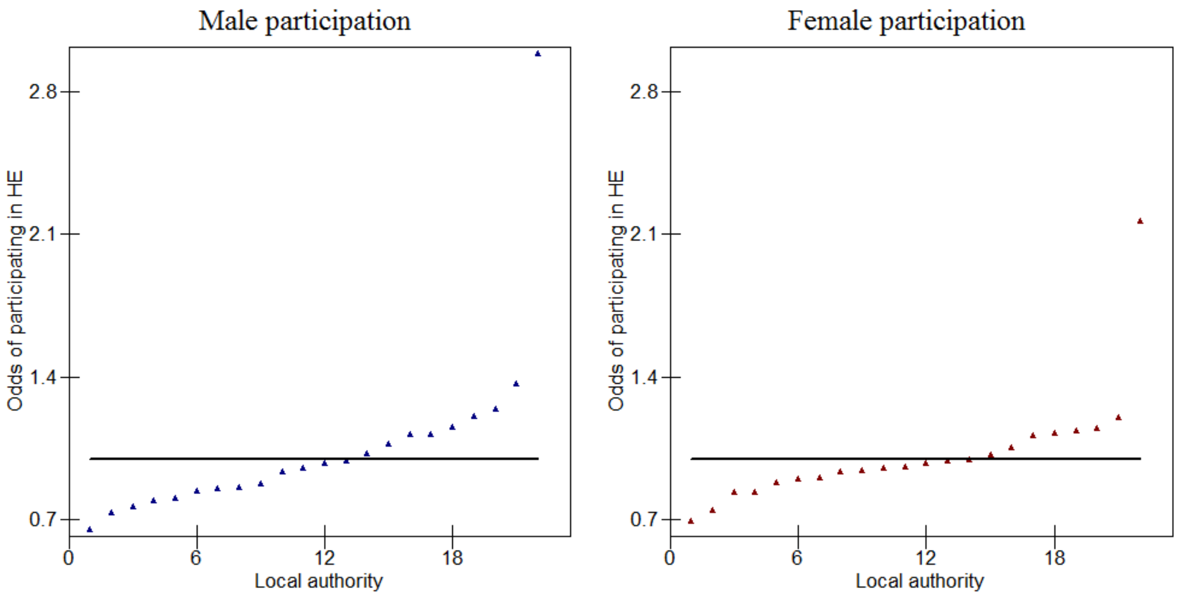


## Local authority effects

One of the potential pitfalls of assessing both school and local authority variance, relating to their hierarchical nature is that significance may be wrongly attributed to one level, which is in fact related to the other level and by estimating the effects separately you may over estimate their effects. One of the advantages of the multilevel approach is that by estimating the effects in one model simultaneously we are able to estimate these effects net of one another and as such can be confident our results are not erroneous.

Figure 2 shows the differences in HE participation according to local authority. These effects are over and above the individual-level effects *and* the school-level effects described in the previous section(s). It is clear that there is substantial between-local authority variance in terms of young people's propensity to participate in HE.

**Figure 2: Shows the odds of participating in HE according to local; authority for males and females**




Whilst HE participation is high in Ceredigion for females and in the Vale of Glamorgan for males, surprisingly, Merthyr Tydfil (shown to have the lowest average GCSE score of all local authorities) has the highest HE participation rates for both males and females. Again, this could be further explored through in-depth qualitative analysis.

Owing to the existence of this very large local authority effect, an additional analysis was conducted using a slightly different outcome variable. Concern was raised, that the local authority effect may relate directly to the form of HE that is predominantly taken in Merthyr Tydfil, namely non-degree level HE. As such it was felt that it would be appropriate that the analysis is run again to measure undergraduate participation in HE. This includes those students registered for Bachelors and Foundation degrees and Certificates of HE, importantly this does not include all 'other' forms of HE, which may fall below Level 4 qualifications. Local authority effects remained statistically significant for males, but disappear for females in relation to undergraduate participation. Importantly, Merthyr Tydfil was no longer shown to have the highest HE participation rates, demonstrating just how important it is to consider carefully what kind of participation we mean.

## Conclusions

This paper set out to (i) replicate Chowdry et al (2013) paper which explores the determinants of participation in HE amongst individuals from socio-economically disadvantaged backgrounds; and (ii) extend this analysis using multilevel modelling and analysis. Consistent with Chowdry et al, this paper has shown that students from the top WIMD quintile are the most likely to participate in HE, with the greatest difference showing at the top of the SES distribution, i.e. between the top and 2$^{nd}$ quintiles. However, unlike in the Chowdry paper, this research found that it is not the bottom quintile, but the 4$^{th}$ quintile who are the least likely group to participate in HE. This might be, but it may be related to the slight differences in the two data sets.

Like Chowdry, it was shown that not all of the socio-economic difference in HE participation arises at the point of entry to HE, but is largely explained by the fact that students living in more deprived areas do not achieve as highly as their more advantaged contemporaries. That said differences between the WIMD quintiles do remain highly significant even after including prior attainment, but are much reduced when compared to the 'raw' estimates.

Ethnicity was shown to be highly pertinent to HE participation. Indeed, it is second only to attainment in terms of the size of its effect. Both 'White other' and 'non-White' groups are significantly more likely to participate than 'White British' students, indeed, non-White students are 12.4% (males) and 10.2% (females) more likely to participate in HE. Only those students categorised as ethnicity unknown are less likely to participate, this was true for both males and females.

It was also shown that schools make a big difference to young people's outcomes however owing to the methodology we are unable to quantify this effect or test any school-level variables to elicit what it is about particular schools that make their students more or less likely to participate in HE.

Although we only have three cohorts we are able to compare their overall participation rates. Male participation has declined year on year between 2005 and 2007. For females however, their participation saw an increase of 1.5% in 2006, but decreased by 0.8% in 2007.

Counter to findings from other work in this field (HEFCE, 2005; Crawford et al., 2010; and DfE, 2010), summer-born students are the most likely to participate in HE, while autumn-born students are the least likely. Whilst this is different to previous research, it is similar to the results from the Chowdry paper.

The adoption of a multilevel approach didn't alter the direction of any of the effects however it did reveal a slightly different relationship between HE participation and the WIMD. The first analysis showed that the fourth WIMD quintile had the lowest HE participation, not, as would be expected, the bottom quintile. However, the multilevel analysis showed the bottom quintile to have the lowest HE participation. It may be that the inclusion of the random effect for local authority has altered this relationship with WIMD.

We can conclude from the multilevel analysis that after taking account of individual differences, schools and local authorities have a significant impact on students' likelihood of participating in HE. This should be of substantial interest to academics and policy makers alike, because it demonstrates that there are equivalently qualified young people whose patterns of HE participation are closely associated with their school and/or local authority, rather than their qualifications. However, these differences associated with schools and local authorities remain unexplained i.e. it is not clear from this analysis what it is about these particular schools and local authorities that make the young people within them more or less likely to participate. It would be hugely beneficial to conduct additional analysis that could further explore this, by considering the effect of particular characteristics about these contexts. It is also noteworthy, that the performance of particular local authorities' changes markedly according to the type of HE participation we are referring to. This highlights the importance of defining what we mean by participation, before we embark on 'widening' it.

## Table 10: Gradients in HE participation for males: results from separate logistic models

| Explanatory Variable | No covariates | + Cohort | + SEN | + FSM | + Ethnicity | + SOB | + GCSE Points | + GCSE CSI | + School |
|---|---|---|---|---|---|---|---|---|---|
| WIMD 2 | -0.540**(0.027) | -0.539**(0.027) | -0.511**(0.028) | -0.519**(0.028) | -0.536**(0.029) | -0.543**(0.029) | -0.255**(0.036) | -0.354**(0.034) | -0.485**(0.030) |
| WIMD 3 | -0.729**(0.028) | -0.730**(0.027) | -0.688**(0.028) | -0.674**(0.028) | -0.729**(0.029) | -0.733**(0.029) | -0.286**(0.035) | -0.458**(0.033) | -0.661**(0.031) |
| WIMD 4 | -1.110**(0.029) | -1.111**(0.029) | -1.066**(0.029) | -1.020**(0.029) | -1.116**(0.029) | -1.114**(0.030) | -0.358**(0.035) | -0.661**(0.034) | -1.077**(0.031) |
| WIMD 5 | -1.586**(0.032) | -1.583**(0.030) | -1.496**(0.031) | -1.397**(0.032) | -1.611**(0.031) | -1.588**(0.032) | -0.443**(0.039) | -0.897**(0.035) | -1.524**(0.034) |
| Year 2006 | | 0.134**(0.023) | | | | | | | |
| year 2007 | | 0.117**(0.023) | | | | | | | |
| sen: Nothing | | | 1.312**(0.041) | | | | | | |
| sen: Action Plus | | | -0.339**(0.079) | | | | | | |
| sen: Statement | | | -0.471**(0.078) | | | | | | |
| FSM: Yes | | | | -0.982**(0.036) | | | | | |
| Eth: White Other | | | | | 0.280**(0.072) | | | | |
| Eth: Non-white | | | | | 0.620**(0.055) | | | | |
| Eth: DK/NS | | | | | -0.169**(0.061) | | | | |
| SOB: Winter | | | | | | -0.017(0.027) | | | |
| SOB: Spring | | | | | | -0.012(0.027) | | | |
| SOB: Summer | | | | | | -0.058*(0.027) | | | |
| GCSE points | | | | | | | 0.098**(0.001) | | |
| GCSE CSI | | | | | | | | 2.540**(0.023) | |
| School | | | | | | | | | 0.261**(0.061) |
| **DIC** | **65033.882** | **64998.419** | **62344.069** | **64210.353** | **64886.781** | **65035.484** | **43296.606** | **50295.773** | **64234.349** |
| **PD** | **4.869** | **6.912** | **7.732** | **5.794** | **7.940** | **8.252** | **5.620** | **5.927** | **112.148** |

***Significance at the 5% level**
****Significance at the 1% level**

## Table 11: Gradients in HE participation for females: results from separate logistic models

| Explanatory Variable | No covariates | + Cohort | + SEN | + FSM | + Ethnicity | + SOB | + GCSE Points | + GCSE CSI | + School |
|---|---|---|---|---|---|---|---|---|---|
| WIMD 2 | -0.450**(0.029) | -0.445**(0.028) | -0.436**(0.028) | -0.425**(0.027) | -0.444**(0.027) | -0.447**(0.029) | -0.245**(0.036) | -0.311**(0.036) | -0.404**(0.029) |
| WIMD 3 | -0.687**(0.029) | -0.685**(0.028) | -0.650**(0.028) | -0.623**(0.028) | -0.683**(0.028) | -0.683**(0.027) | -0.297**(0.037) | -0.431**(0.034) | -0.633**(0.029) |
| WIMD 4 | -1.139**(0.030) | -1.137**(0.029) | -1.105**(0.030) | -1.030**(0.029) | -1.139**(0.030) | -1.136**(0.029) | -0.441**(0.036) | -0.719**(0.035) | -1.111**(0.030) |
| WIMD 5 | -1.592**(0.031) | -1.589**(0.031) | -1.533**(0.030) | -1.370**(0.031) | -1.611**(0.030) | -1.589**(0.029) | -0.443**(0.039) | -0.939**(0.036) | -1.545**(0.032) |
| Year 2006 | | 0.208**(0.022) | | | | | | | |
| year 2007 | | 0.172**(0.022) | | | | | | | |
| sen: Nothing | | | 1.470**(0.048) | | | | | | |
| sen: Action Plus | | | -0.119(0.095) | | | | | | |
| sen: Statement | | | -0.681**(0.130) | | | | | | |
| FSM: Yes | | | | -1.083**(0.034) | | | | | |
| Eth: White Other | | | | | 0.365**(0.074) | | | | |
| Eth: Non-white | | | | | 0.549**(0.053) | | | | |
| Eth: DK/NS | | | | | -0.143*(0.061) | | | | |
| SOB: Winter | | | | | | -0.013(0.027) | | | |
| SOB: Spring | | | | | | -0.044(0.026) | | | |
| SOB: Summer | | | | | | -0.058*(0.026) | | | |
| GCSE points | | | | | | | 0.101**(0.001) | | |
| GCSE CSI | | | | | | | | 2.488**(0.022) | |
| School | | | | | | | | | 0.136**(0.028) |
| **DIC** | **68483.737** | **68390.134** | **66458.222** | **67289.628** | **68353.353** | **68482.918** | **46021.136** | **53192.700** | **67627.609** |
| **PD** | **5.189** | **6.893** | **7.936** | **5.953** | **8.089** | **8.041** | **6.047** | **6.104** | **104.485** |

***Significance at the 5% level**
****Significance at the 1% level**

## Table 12: Gradients in HE participation for males: results from separate logistic models

| Explanatory Variable | No covariates | + School | + LEA | Random Effects | Random Effects + GCSE | Random Effects + SEN | Random Effects + FSM | Random Effects + ETH | Random Effects + SOB | Random Effects + COHORT |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed part** | | | | | | | | | | |
| WIMD 2 | -0.540*** | -0.485*** | -0.563*** | -0.456*** | -0.130*** | -0.432*** | -0.437*** | -0.458*** | -0.454*** | -0.456*** |
| WIMD 3 | -0.729*** | -0.661*** | -0.786*** | -0.623*** | -0.165*** | -0.593*** | -0.579*** | -0.629*** | -0.623*** | -0.625*** |
| WIMD 4 | -1.110*** | -1.077*** | -1.180*** | -0.992*** | -0.293*** | -0.936*** | -0.921*** | -1.003*** | -0.994*** | -0.996*** |
| WIMD 5 | -1.586*** | -1.524*** | -1.683*** | -1.409*** | -0.408*** | -1.321*** | -1.255*** | -1.430*** | -1.412*** | -1.411*** |
| GCSE Points | | | | | 0.106*** | | | | | |
| SEN: Nothing | | | | | | 1.389*** | | | | |
| SEN: AP | | | | | | -0.316*** | | | | |
| SEN: Statement | | | | | | -0.345*** | | | | |
| FSM: Yes | | | | | | | -0.956*** | | | |
| Ethnicity: White Other | | | | | | | | 0.311*** | | |
| Ethnicity: Non-White | | | | | | | | 0.717*** | | |
| Ethnicity: DK/NS | | | | | | | | -0.218*** | | |
| Winter | | | | | | | | | -0.031 | |
| Spring | | | | | | | | | -0.018 | |
| Summer | | | | | | | | | -0.066** | |
| 2006 | | | | | | | | | | 0.135*** |
| 2007 | | | | | | | | | | 0.122*** |
| **Random part** | | | | | | | | | | |
| School | | 0.261*** | | 0.217*** | 0.206*** | 0.149*** | 0.176*** | 0.221*** | 0.215*** | 0.215*** |
| LEA | | | 0.065*** | 0.024 | 0.121** | 0.049** | 0.036* | 0.019 | 0.022 | 0.020 |
| **DIC** | **65033.882** | **64234.349** | **64608.596** | **63487.802** | **41773.020** | **60901.895** | **62780.307** | **63326.460** | **63489.065** | **63455.072** |
| **Diff in DIC** | | -799.533 | -425.286 | -1546.08 | -23,260.862 | -4131.987 | -2253.575 | -1707.422 | 1544.817 | 1578.81 |
| **PD** | **4.869** | **112.148** | **25.036** | **216.103** | **195.862** | **203.774** | **210.251** | **218.899** | **218.863** | **216.715** |

***Significance at the 10% level**
**** Significance at the 5% level**
*****Significance at the 1% level**

## Table 13: Gradients in HE participation for females: results from separate logistic models

| Explanatory Variable | No covariates | + School | + LEA | Random Effects | Random Effects + GCSE | Random Effects + SEN | Random Effects + FSM | Random Effects + ETH | Random Effects + SOB | Random Effects + COHORT |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed Part** | | | | | | | | | | |
| WIMD 2 | -0.450*** | -0.404*** | -0.515*** | -0.392*** | -0.159*** | -0.384*** | -0.374*** | -0.395*** | -0.394*** | -0.395*** |
| WIMD 3 | -0.687*** | -0.633*** | -0.801*** | -0.658*** | -0.232*** | -0.634*** | -0.608*** | -0.663*** | -0.659*** | -0.665*** |
| WIMD 4 | -1.139*** | -1.111*** | -1.244*** | -1.055*** | -0.428*** | -1.019*** | -0.969*** | -1.064*** | -1.059*** | -1.063*** |
| WIMD 5 | -1.592*** | -1.545*** | -1.700*** | -1.425*** | -0.453*** | -1.370*** | -1.244*** | -1.444*** | -1.429*** | -1.431*** |
| GCSE Points GM | | | | | 0.107*** | | | | | |
| SEN: Nothing | | | | | | 1.536*** | | | | |
| SEN: AP | | | | | | -0.061 | | | | |
| SEN: Statement | | | | | | -0.578*** | | | | |
| FSM: Yes | | | | | | | -1.046*** | | | |
| Ethnicity: White Other | | | | | | | | 0.327*** | | |
| Ethnicity: Non-White | | | | | | | | 0.690*** | | |
| Ethnicity: DK/NS | | | | | | | | -0.216*** | | |
| Winter | | | | | | | | | -0.004 | |
| Spring | | | | | | | | | -0.039 | |
| Summer | | | | | | | | | -0.053** | |
| 2006 | | | | | | | | | | 0.206*** |
| 2007 | | | | | | | | | | 0.175*** |
| **Random part** | | | | | | | | | | |
| School | | 0.136*** | | 0.231*** | 0.202*** | 0.193*** | 0.199*** | 0.231*** | 0.233*** | 0.230*** |
| LEA | | | 0.073*** | 0.035** | 0.073** | 0.053* | 0.043* | 0.039* | 0.035 | 0.037 |
| **DIC** | **68483.737** | **67627.609** | **67981.123** | **66399.172** | **44591.697** | **64481.865** | **65374.498** | **66237.676** | **66395.334** | **66308.706** |
| **Difference in DIC** | | 856.128 | 502.614 | 2084.565 | 23892.04 | 4001.872 | 3109.239 | 2246.061 | 2088.403 | 2175.031 |
| **PD** | **5.189** | **104.485** | **25.118** | **216.895** | **197.015** | **210.761** | **211.500** | **217.757** | **218.556** | **217.890** |

***Significance at the 10% level**
**** Significance at the 5% level**
*****Significance at the 1% level**

## Table 14: Gradients in HE participation for Males: results from sequential[3] logistic models

| Explanatory Variable | Random Effects + GCSE | Previous Model + SEN | Previous Model + FSM | Previous Model + ETH | Previous Model + COHORT | Previous Model + SOB |
|---|---|---|---|---|---|---|
| **Fixed Part** | | | | | | |
| WIMD 2 | -0.130*** | -0.128*** | -0.129*** | -0.128*** | -0.139*** | -0.128*** |
| WIMD 3 | -0.165*** | -0.159*** | -0.157*** | -0.161*** | -0.170*** | -0.163*** |
| WIMD 4 | -0.293*** | -0.289*** | -0.287*** | -0.293*** | -0.304*** | -0.292*** |
| WIMD 5 | -0.408*** | -0.408*** | -0.397*** | -0.415*** | -0.423*** | -0.412*** |
| GCSE Points GM | 0.106*** | 0.108*** | 0.108*** | 0.108*** | 0.108*** | 0.108*** |
| SEN: Nothing | | -0.196*** | -0.188*** | -0.183*** | -0.216*** | -0.178*** |
| SEN: AP | | 0.002 | 0.010 | 0.026 | -0.008 | 0.023 |
| SEN: Statement | | 0.611*** | 0.627*** | 0.647*** | 0.613*** | 0.647*** |
| FSM: Yes | | | -0.098** | -0.128*** | -0.128*** | -0.129*** |
| Ethnicity: White Other | | | | 0.285** | 0.288** | 0.294** |
| Ethnicity: Non-White | | | | 0.856*** | 0.862*** | 0.867*** |
| Ethnicity: DK/NS | | | | 0.004 | -0.001 | 0.012 |
| 2006 | | | | | 0.028 | 0.030 |
| 2007 | | | | | -0.021 | -0.014 |
| Winter | | | | | | 0.110*** |
| Spring | | | | | | 0.197*** |
| Summer | | | | | | 0.243*** |
| **Random Part** | | | | | | |
| School | 0.206*** | 0.207*** | 0.205*** | 0.202*** | 0.198*** | 0.204*** |
| LEA | 0.121** | 0.135** | 0.141** | 0.135** | 0.140** | 0.129** |
| **DIC** | **41773.020** | **41696.101** | **41692.485** | **41576.329** | **41575.995** | **41524.379** |
| **Difference in DIC** | | -76.919 | -3.616 | -116.156 | -0.334 | -51.616 |
| **PD** | **195.862** | **199.469** | **199.206** | **202.814** | **203.528** | **206.275** |

***Significance at the 10% level**
**** Significance at the 5% level**
*****Significance at the 1% level**

[3] The ordering of the sequence of models is determined by the size of the reduction in the DIC from the separate modelling i.e. the explanatory variables with the most predictive power, will be added first.

## Table 15: Gradients in HE participation for Females: results from sequential logistic models

| Explanatory Variable | Random Effects + GCSE | Previous Model + SEN | Previous Model + FSM | Previous Model + ETH | Previous Model + COHORT | Previous Model + SOB |
|---|---|---|---|---|---|---|
| **Fixed Part** | | | | | | |
| WIMD 2 | -0.159*** | -0.158*** | -0.159*** | -0.162*** | -0.159*** | -0.157*** |
| WIMD 3 | -0.232*** | -0.230*** | -0.225*** | -0.229*** | -0.230*** | -0.225*** |
| WIMD 4 | -0.428*** | -0.425*** | -0.414*** | -0.422*** | -0.423*** | -0.416*** |
| WIMD 5 | -0.453*** | -0.450*** | -0.421*** | -0.438*** | -0.441*** | -0.436*** |
| GCSE Points GM | 0.107*** | 0.108*** | 0.107*** | 0.107*** | 0.107*** | 0.108*** |
| SEN: Nothing | | -0.096 | -0.066 | -0.047 | -0.092 | -0.072 |
| SEN: AP | | 0.081 | 0.119 | 0.123 | 0.082 | 0.091 |
| SEN: Statement | | 0.413 | 0.457*** | 0.477*** | 0.426*** | 0.430*** |
| FSM: Yes | | | -0.219*** | -0.242*** | -0.242*** | -0.236*** |
| Ethnicity: White Other | | | | 0.325*** | 0.343*** | 0.344*** |
| Ethnicity: Non-White | | | | 0.699*** | 0.703*** | 0.694*** |
| Ethnicity: DK/NS | | | | -0.099 | -0.088 | -0.092 |
| 2006 | | | | | 0.152*** | 0.153*** |
| 2007 | | | | | 0.106*** | 0.110*** |
| Winter | | | | | | 0.110*** |
| Spring | | | | | | 0.130*** |
| Summer | | | | | | 0.223*** |
| **Random Part** | | | | | | |
| School | 0.202*** | 0.204*** | 0.203*** | 0.199*** | 0.199*** | 0.197*** |
| LEA | 0.073** | 0.072* | 0.072* | 0.073* | 0.072* | 0.076* |
| **DIC** | **44591.697** | **44582.387** | **44558.742** | **44471.279** | **44446.283** | **44405.958** |
| **Difference in DIC** | | -9.31 | -23.645 | -87.463 | -24.996 | -40.325 |
| **PD** | **197.015** | **199.999** | **201.335** | **204.793** | **206.250** | **208.734** |

***Significance at the 10% level**
**** Significance at the 5% level**
*****Significance at the 1% level**

## References

Bukodi, E. and Goldthorpe, J. (2013) Decomposing "Social Origins": the effects of parents' class, status, and education on the educational attainment of their children. *European Sociological Review*, 29, (5), p.1024-1039.

Chowdry, H. et al. (2013) Widening Participation in Higher Education: Analysis using Linked Administrative Data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 176, (2), p.431–457

Crawford et al. (2010) When you are born matters: the impact of date of birth on educational outcomes in England

DfE (2010) Month of Birth and Education. London: Department for Education.

HEFCE (2005) *Young participation in higher education 1994-2000.* Bristol: HEFCE.

Hobbs, G. & Vignoles, A. (2010) Is children's free school meal 'eligibility' a good proxy for family income? *British Educational Research Journal*, 36**,** p.673-690.

Jones, K. and Subramanian, S. (2012) Developing multilevel models for analysing contextuality, heterogeneity and change using MLwiN 2.2 Volume 1. Bristol: University of Bristol.

Robinson, W. (1950) Ecological Correlations and the Behavior of Individuals. *American Sociological Review*, 15, (3), p.351-357.

Taylor, C. et al. (2013) Creating an inclusive Higher Education system? Progression and outcomes of students from low participation neighbourhoods at a Welsh university. *Contemporary Wales*, 26, p.138-161.