Understanding KS4 attainment and progress: evidence from LSYPE2

Research report

October 2018

Carli Lessof, Andy Ross, Richard Brind, Carrie Harding, Emily Bell and George Kyriakopoulos - Kantar Public

Acknowledgements

Contents

List of figures

List of tables

List of boxes

Executive Summary

Report overview

This research report, based on data from the Longitudinal Study of Young People in England Cohort 2, focuses on the academic attainment of young people in year 11. This is the point at which most sit their GCSE exams (or equivalents), which are also referred to as key stage4 (or KS4) qualifications.

The report addresses three main questions:

- What are the 'face value' differences in the attainment of different groups?

  - Here we look at the absolute differences between different groups (e.g. do children with degree educated mothers typically achieve higher KS4 scores than those whose mothers are less highly educated?)

- What are the underlying factors that predict attainment?

  - While there may be 'face value' differences in attainment between different groups, these are often driven by other underlying factors (e.g. some 'face value' differences associated with ethnicity may actually be driven by differences in socioeconomic circumstances and parental expectations rather than by ethnicity per se)

  - As such, we employ a statistical model which allows us to understand which variables predict attainment, and to what extent, after controlling for all the other factors in the model

- Which factors predict progress at secondary school?

  - The attainment of young people at KS4 is affected not only by the period during secondary school, but also by influences that came before, including their primary and pre-school development

  - We therefore employ a second statistical model, which further adjusts for prior attainment using key stage2 (KS2) scores – these are the assessments taken in the final year of primary school. Consequently, this model tells us the extent to which different factors predict the progress young people make during secondary school.

It should be noted that, while we identify factors which statistically predict attainment and progress, the associations that we identify do not allow us to infer causality. A finding might support a particular hypothesis but it is not its absolute proof.

All of the analyses in the report are based on a sample drawn from the cohort of young people in England who completed year 11 in 2015.

Key findings

Absolute differences in attainment at KS4

There is already much published data on absolute differences in attainment at KS4. This has shown, for example, that attainment is higher for girls and lower for those eligible for free school meals[1]. Some of the absolute differences in attainment which are not available in the official GCSE (and equivalent) results data include:

- In absolute terms (i.e. not taking account of other factors such as deprivation), the attainment of young people living with two biological parents was higher, on average, than that of young people living with one biological parent and markedly higher than that of young people living with neither of their biological parents

- There was a relatively linear relationship between the level of maternal qualification and the young person's attainment, with higher levels of maternal qualification associated with higher attainment for their children

- Young people living in council and housing authority rental properties had lower attainment, on average, than those living in private rental properties who, in turn, had lower attainment than those in owned or mortgaged properties

The underlying factors which predict attainment at KS4

Moving on to examine findings from the statistical modelling, a broad range of factors were found to be predictive of attainment. Some of the key findings are grouped into broad themes below.

Disadvantage

Many aspects of disadvantage were associated with lower attainment, even after controlling for other factors. And, just as importantly, disadvantage is likely to have a layering effect – the more elements of disadvantage a young person faces, the lower their attainment is likely to be.

Elements of disadvantage which were found to be associated with lower attainment, even after controlling for all the other modelled factors, include:

- the young person living with neither biological parent

- living in rented accommodation

- living in a household with below median income

- having a mother without Higher Education or degree level qualifications

- absence of internet-connected home desktops or laptops

---

[1] See: https://www.gov.uk/government/statistics/revised-gcse-and-equivalent-results-in-england-2014-to-2015

- living in more disadvantaged areas

Health and wellbeing

Several factors relating to health and wellbeing, which in many cases effectively represent further types of disadvantage, were also found to be predictive of attainment. In particular:

- Psychological distress showed a curvilinear relationship with attainment – young people reporting particularly low levels of psychological distress or particularly high levels of psychological distress had lower attainment than those with moderate levels of psychological distress

- Young people who suffered from cyber-bullying, all other factors being equal, were found to have lower attainment than those who did not

- Young people with challenging parental relationships showed lower attainment regardless of their background – those whose parents reported that they argued more than once a week had lower attainment than those who argued less often

- Sleeping for nine and a half hours or more per night (that is, more than the optimal amount of sleep) was associated with lower attainment (even after factors such as psychological distress and the other modelled variables had been controlled for)

Approach to school

Factors relating to the approach of young people to their schooling were generally found to relate to attainment in ways which might intuitively be expected:

- Truanting for several days at a time (or longer) corresponded to markedly lower attainment
- Young people who usually did all of their homework had higher attainment than those who usually did most of their homework and, particularly, those who usually only did some or none of it

The gender difference

While the attainment of boys was lower than that of girls in absolute terms, this gap was no longer significant when other factors were controlled for. Related factors, which served to explain the difference in gender attainment, included learning difficulties and factors relating to the home environment, for example:

- The higher incidence of Special Educational Needs (SEN) and disabilities which affect schooling among boys

- The fact that parents were more likely to *expect* their daughters to go on to university and were less likely to *want* their daughters to pursue paths outside education after year 11 (e.g. apprenticeships)

Ethnicity

It is also worth noting that after controlling for all the other modelled factors, the attainment of young people from Black African and Black Caribbean backgrounds was slightly lower than that of other groups, a finding that would merit further investigation.

The predictors of progress at secondary school

Finally, we examine the factors which correspond with progress in attainment at secondary school. It is interesting to note that while some predictors of progress are the same as the predictors of attainment, this is not always the case.

Disadvantage

Elements of disadvantage, as well as being predictive of lower attainment overall, were also found to be predictive of lower progress throughout secondary school, suggesting that there is a widening attainment deficit as disadvantaged young people move from year 7 to year 11.

However, though their attainment was lower, young people in schools with a higher proportion of FSM eligible students progressed slightly <u>more</u> than those in schools with a lower proportion, all other modelled factors being equal.

Pupils with SEN made less progress (in terms of GCSE or equivalent points) during secondary school than those without SEN: this difference would be expected given that those with SEN include young people with moderate, severe and profound learning difficulties and/or other types of needs and behaviours that challenge their educational attainment.

Negative behaviours

Some of the school behaviours of young people were found to correspond negatively with progress. In particular:

- Truanting for several days at a time (or longer), as well as corresponding to lower attainment, also corresponded to lower progress

- Although misbehaving in class did not predict differences in attainment overall, it did predict school progress - those who reported misbehaving in half or more classes made less progress on average than those who did not

Catching up

Young people born in the summer term (i.e. those who were youngest in their school year group) made more progress in the first five years of secondary school than those

born in the autumn term. They were effectively catching up, to at least some extent, from a lower KS2 starting point.

Ethnicity

There was not a statistically significant difference in the amount of progress that young people from Black African and Black Caribbean backgrounds made throughout secondary school compared to White students. This suggests that the attainment deficit at KS4 (discussed above) is more attributable to the primary or pre-school phases and there was no catch up during secondary school.

Implications

Some of our findings support other established evidence – for example, the relationship between lower attainment and lower household income, not living with parents, poor health, learning difficulties and poor psychological wellbeing.

These material, social and individual challenges are well-recognised types of disadvantage for which resources and services are already providing support. Nevertheless, the fact that these issues continue to predict attainment deficits underlines the importance of ongoing efforts to address them.

Our other findings relate to less well-recognised or less robustly quantified factors which predict attainment – these span diverse topics such as the influence of the home environment; prosocial and negative behaviours; levels of psychological distress; and relationships with teachers.

These have led us to a broad range of hypotheses and to highlight research questions which could usefully be investigated further to inform possible changes to policy or practice. We discuss this more fully in our conclusions section (see Chapter 5) but examples include:

- Why are parents more inclined to encourage girls to pursue academic pathways rather than boys? Would making them aware of this bias and its effects improve boys' attainment?
- Given the link between attainment and maternal qualifications, would improving mothers' qualification levels during their child's lifetime improve the young person's own attainment?
- Why does a lack of a home computer connected to the internet lead to lower attainment? Where is this most problematic and could this issue be addressed by government and/or charities?

Of all the findings discussed in this report, perhaps the most striking is the apparent layering of disadvantage. Our analysis suggests that the more elements of disadvantage that a young person faces, the lower their attainment will typically be, potentially entrenching that disadvantage into their later life.

This is a complex topic which spans the remit of government departments and public services. As such, one of the underlying themes of the report is that cross governmental action will be needed to address the social problems causing attainment deficits – they are not purely an educational issue.

Methodology

LSYPE 2 began in 2012/13 and tracks a sample of 13,100 young people in England from the age of 13/14 annually for seven years, through to the age of 19/20. The data in this report are primarily drawn from the second wave of LSYPE2, with supplementary information drawn from the National Pupil Database and other sources of administrative data where applicable (e.g. the IDACI measure of area deprivation). These data pre-date the introduction of the new GCSE grading system and our analyses are therefore based on the legacy A*-G scale[2].

This report is based only on young people in schools in the maintained sector. Young people who attended special schools and young people for whom the necessary administrative data was missing, or whose LSYPE2 data included missing values, were excluded from the statistical models. Data were weighted to take account of attrition, and an assessment of the impact of missing data suggests estimations of the extent to which disadvantage corresponds with lower attainment are likely to be conservative.

While we discuss the extent to which different factors correspond with attainment, it is important to note that we cannot infer causality. To do that, we would need clear evidence about how the causal relationship works, which is not something we can easily do with an observational study of this kind or with this statistical approach. As such, a finding might support a particular hypothesis but it is not its absolute proof.

In our base model 84% of the variation in pupils' attainment was attributed to differences between pupils, and 16% was attributed to differences between the schools they attended. Although, it is important to note that the differences in attainment attributed to schools is also related to differences in the kinds of pupils attending those schools. Using our models, we were able to explain 57% of the variation in attainment attributed to the individual and 87% attributed to schools.

Further details of the methodology can be found in Chapter 1 and in the accompanying technical report.

---

[2] Interviews were conducted in 2015, at which point the A*-G grading system was still in place. A new grading system, ranging from 1-9, is currently being introduced through a phased roll-out across 2017 and 2018. Further details can be found here:
https://www.gov.uk/government/news/new-gcse-9-to-1-grades-coming-soon

Chapter 1 Introduction

Aims of the report and its relevance to policy

This research report focuses on the academic attainment of young people in year 11, the point at which most sit their GCSE exams (or equivalent), also referred to as key stage 4 (or KS4) qualifications. Attainment at this stage of the lives of young people can have profound impacts on their subsequent path through life. Although the results which young people attain at KS4 are not absolute determinants of later outcomes, they can have direct impacts on their journey through the next stages of education and training, and also on their later working life[3].

All of our analyses are based on the cohort of young people in England who completed year 11 in 2015, prior to the introduction of the new GCSE grading system[4]. As the report progresses, we look to understand what lies behind differences in attainment - sometimes factors which appear significant at face value are less important than they may at first seem or are operating in more complex ways than is initially apparent.

To begin, in Chapter 2, we examine the absolute differences in attainment between different groups of young people, for example those whose mothers hold different levels of qualifications or those with different types of housing tenure. This initial analysis provides us with a solid starting point, allowing us to understand, to cite just one illustrative example, whether the absolute attainment of students with university educated mothers differs from those whose mothers hold no formal qualifications.

However, in order to understand these differences more fully, it is necessary to consider the ways in which many measures may interrelate. Using another example, if we find that there is a performance difference on average between students from White and Black Caribbean ethnic backgrounds, rather than assume that this is an effect which relates directly to their ethnicity, we explore whether it is, at least in part, a product of other differences between these groups, for example their average level of household income or home environment. As such, in Chapter 3, we create a statistical model which takes account of a very broad range of measures – approaching ninety - simultaneously. This enables us to better understand the key predictors of attainment at KS4.

Finally, moving beyond the focus on attainment, we also investigate the amount of progress that young people make through secondary school. To do this, in Chapter 4, we

---

[3] Earlier analysis of LSYPE data has, for example, identified a relationship between low attainment at GCSE and NEET (Not in Employment, Education or Training) status at the age of 18/19.
See: Young people's education and labour market choices aged 16/17 to 18/19, Crawford et al, 2011
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/183355/DFE-RR182.pdf
[4] Interviews were conducted in 2015, at which point the A*-G grading system was still in place. A new grading system, ranging from 1-9, is currently being introduced through a phased roll-out across 2017 and 2018. Further details can be found here:
https://www.gov.uk/government/news/new-gcse-9-to-1-grades-coming-soon

take account of key stage 2 (KS2) performance – the attainment measure collected towards the end of primary education – by adding it to our model. This enables us to better understand, for example, whether young people with low attainment at KS4 have made less progress than their peers throughout secondary school, or whether the performance deficit actually arises from earlier in their lives[5].

By improving our understanding of these relationships, it will be possible to better identify the factors which correspond with attainment and to understand where inequalities lie, with a view to shaping policy to reduce those inequalities and increase social mobility. In doing so, there is also the potential to raise standards overall, thereby improving the standing of young people living in England in a highly competitive and connected world.

From an education policy perspective, as well as the perspective of the young people themselves, attainment at KS4 is highly important. KS4 results represent one of the key indicators used by employers to select candidates for jobs[6]. As such, if certain groups underperform and this underperformance can be meaningfully addressed, this will have a beneficial effect in terms of addressing enduring social division, one of the five planks of the government's manifesto[7].

This is particularly important, given the essentially sequential nature of the education system whereby progress to the next stage of education is typically dependent on attainment at the previous stage. The ability to move on to study for A levels or higher/advanced apprenticeships can depend on KS4 attainment. In turn, the most straightforward path to university requires the achievement of these post-KS4 qualifications. As such, under-attainment at KS4 effectively limits the opportunities of the young person to progress in their education, which in turn impacts on their life chances[8]. Again, bearing in mind the pledge to address enduring social division, this is a vital issue to address – essentially speaking, the inequalities in KS4 attainment which often arise from disadvantage, as discussed throughout the report, will continue to ripple well past year 11. Maths and English are particularly important in this respect as they are often included as a pre-requisite for admission to Further Education courses and a range of

---

[5] Early life factors such as attendance or non-attendance at pre-school have been found to have ongoing impacts on attainment at KS4, as highlighted in the EPPSE 3-16+ study, Sylva et al, 2014. See: http://www.ucl.ac.uk/ioe/research/pdf/16-educational-Developmental-Outcomes-RR.pdf

[6] For example, research conducted by Ofqual found that only 6% of employers said that they did *not* consider GCSE grades when reviewing applications from candidates.
See: New GCSE Grades Research Amongst Employers, Wiseman and Parry, 2013
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/529390/2013-11-01-bmg-research-with-employers-on-new-gcse-grades.pdf

[7] The government's manifesto can be found here:
https://www.conservatives.com/manifesto

[8] For example, research by BIS identified a broad range of benefits associated with attendance at university, spanning topics including (but not limited to) mental health, risky behaviours (e.g. excessive drinking and smoking), civic engagement and income. See:
The Benefits of Higher Education Participation for Individuals and Society: key findings and reports "The Quadrants", 2013.
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/254101/bis-13-1268-benefits-of-higher-education-participation-the-quadrants.pdf

careers. To this end, we include analyses focussing specifically on these two subjects, along with broader measures of attainment.

Policy initiatives

In recent years there have been many important policy developments which are relevant to attainment at KS4. The following is by no means a complete discussion of these policy changes, but does provide an overview of those factors which we believe to be most relevant for this cohort.

These policy developments include:

- Raising the participation age to ensure young people remain in education or training for longer
- A greater focus on vocational qualifications and apprenticeships
- Changes to GCSEs and A levels
- Introduction of the English Baccalaureate (EBacc)
- Significant increases in university tuition fees
- The expansion of academy schools and introduction of free schools

One of the most significant policy developments for our cohort has been the **increase in the participation age**. Whereas earlier cohorts were able to leave the education and training system entirely at the age of 16, the cohort of young people who were in year 11 in 2015 were required to stay in education or training until the age of 18, with a view to boosting skills and reducing the number of young people not in education, employment or training (NEET). There has also been an increasing emphasis on the validity of apprenticeships as an alternative to non-vocational qualifications, with the number of apprenticeship starts in England increasing from 189,000 in the 2004/5 academic year to 499,900 in 2014/15, though the rate of increase amongst those aged 18 or younger has, thus far, been less rapid (rising from 113,520 in 2004/5 to 125,850 in 2014/15) [9]. As such, year 11 students in 2015 have been growing up in an environment in which they are expected to remain in education for longer and to achieve a higher final level of qualification than their predecessors.

Those who would otherwise have chosen to leave the education or training system entirely, therefore, face a very different path through their late-teen years than their predecessors.

There have also been significant changes to GCSE and A level syllabuses with perhaps the most significant development (from the perspective of year 11 students in 2014/15) being **the move from modular to linear GCSEs** in 2012. In basic terms, this change

---

[9] Data sourced from FE data library:
Apprenticeships by geography, equality & diversity and sector subject area: starts 2002/03 to Q1 2016/17
https://www.gov.uk/government/statistical-data-sets/fe-data-library-apprenticeships

means that there is now a greater focus on assessments at the end of the GCSE courses and the option to re-sit modules mid-course has been removed. As mentioned earlier, it should be noted that the young people in the LSYPE2 cohort were not (in year 11 at least) directly affected by the changes to the GCSE grading system which began in 2017.

Another significant policy development has been the introduction of the **English Baccalaureate (EBacc)** in 2010. This was designed as a measure of KS4 performance and, at the time of our interviews in 2015, was based on the proportion of students in a school achieving a Grade A*-C in five 'traditional' GCSE subjects including: English; maths; science; history or geography; and a foreign language. As such, there may have been a shift in the emphasis that schools placed on these subjects for the LSYPE2 cohort and evidence suggests that pupils at schools which were early adopters of the EBacc largely benefitted from the changes in curriculum[10]. The Progress 8 and Attainment 8 measures (which take account of a broader range of subjects) had been announced but had not yet been implemented at the time the LSYPE2 cohort were in year 11. As such, they should not have had a major impact on their education.

Looking to later life stages, which may nevertheless impact on aspirations and attainment at year 11, the **increase in the cap on tuition fees** that universities are permitted to charge is another potentially significant factor. In 2005 the maximum permissible tuition fee was £3,000 per year and by 2015 this had risen to £9,000.

Another point worth noting is that the **number of academy schools has increased dramatically** in recent years, with rapid change occurring from April 2011[11]. This change means that many young people in 2015 were in an environment where they did not necessarily follow the full national curriculum. The circumstances of schools becoming academies were also initially skewed to either 'Outstanding' schools or those which were significantly under-performing. Because this period of peak conversion to academy status overlaps with the LSYPE2 cohort journey through secondary school, we have not included school type in our analyses (in basic terms, the data would be unclear as many young people would have attended a school which did not have a consistent type throughout their time at it).

Alongside these policy developments, it is also worth bearing in mind the wider societal changes experienced by the young people in the LSYPE2 cohort. These have been widely discussed elsewhere (including in previous reports based on LSYPE2 data), so we do not go over this ground again here.

---

[10] This is discussed further in the 'Changing the Subject' report from the Sutton Trust (Thompson and Allen, 2016). See:
https://www.suttontrust.com/research-paper/changing-the-subject/
[11] For example, between April 2011 and April 2014, the number of academies increased from 465 to 4,010 (of which 200 were free schools). Data sourced from Impact indicators 23, 24, 25: number of academies and free schools as a percentage of the total state-funded mainstream schools, March 2015 update

Trends in attainment

National statistics showing trends in KS4 attainment are published annually[12]. We do not describe these at length here, but key findings from the 2015 cohort were:

- 53.8% of pupils achieved 5 or more GCSEs (or equivalent) at A*-C

    o This marked an increase from 53.4% in 2014, but remained below the levels recorded from 2011 (59.0%) to 2013 (59.2%)

- The performance gap between boys and girls narrowed in terms of the proportion attaining 5 or more GCSEs (or equivalent) at A*-C

    o In 2014 the proportion of boys reaching this standard was 10.7% lower than the proportion of girls. In 2015 this difference had fallen to 9.9%

- The attainment gap between disadvantaged pupils and others (based on the gap index) increased by 1.6% in 2015. The gap was, nevertheless, 6.6% smaller than that observed in 2011.

Methodology

A more detailed summary of the LSYPE2 methodology can be found in Appendix A, with further detail published in the Technical Reports[13]. As such, we only present a brief summary of the key points here:

- The survey data for this report are primarily drawn from the second wave of LSYPE2 which was conducted in the 2013/14 academic year, when the young people were in year 10

- The LSYPE2 survey data were collected through face to face interviews with both the young people themselves and their parents/guardians

- All data in this report are based on young people attending maintained schools

- The survey data are supplemented with linked data which provides valuable additional information. Examples of this linked data include: aggregate information about the school each young person attends; characteristics about the area where they live; and their attainment at KS2 and KS4

- Each chapter in this report addresses a different theme relating to attainment and each employs a different analytical approach

---

[12] Those corresponding to the 2015 cohort are outlined in the National Statistic 'Revised GCSE and equivalent results in England: 2014 to 2015', 21st January 2016. See:
https://www.gov.uk/government/statistics/revised-gcse-and-equivalent-results-in-england-2014-to-2015
[13] Published here:
http://doc.ukdataservice.ac.uk/doc/7813/mrdoc/pdf/7813technical_report_wave_1.pdf and
http://doc.ukdataservice.ac.uk/doc/7813/mrdoc/pdf/7813technical_report_lsype_wave_2_and_3_v6.pdf

- o Chapter 2 employs bivariate analysis to determine the absolute differences in attainment between different groups of young people

- o Chapter 3 employs multi-level multivariate regression modelling to better understand how specific factors correspond to attainment after controlling for other factors (e.g. to what extent does ethnicity correspond with attainment after controlling for factors such as socioeconomic circumstances and parental expectations?)

- o Chapter 4 further adjusts for young people's attainment at KS2[14] – to examine which of the measures already considered best explain *progress* in attainment between KS2 and KS4 after controlling for other factors

- Each chapter has a different total sample size, primarily reflecting the fact that the analyses in later chapters require records where there are valid answers to all measures included in the models. As such, Chapter 2 is based on a starting sample of 9,350 individuals, Chapter 3 is based on 4,200 individuals and Chapter 4 is based on 4,091 individuals.

Limitations

There are a number of potential limitations to bear in mind when interpreting the findings outlined in this report, though it should equally be noted that LSYPE2 represents a robust data source and the analyses have been conducted with an emphasis on methodological rigour. Again, we present a brief summary here and a more detailed discussion of these points can be found in Appendix A.

- While LSYPE2 and the linked data used in this report represent a rich source of information, it is not possible to include every potential influencer of attainment in the statistical models

- It was necessary to base the statistical models on a sub-sample of young people for whom all data was available – the primary effect of this is that we are likely to understate the strength of the relationship between disadvantage and attainment. The sub-sample included in the models also had slightly higher and less dispersed attainment scores than were found in the total sample. An alternative approach which could have been used and would have retained a larger sample size involves multiple imputation of missing data but was not possible given the resource available for this project

- That said, we were able to include imputed KS2 scores for many young people whose data would otherwise have been missing, primarily because two teachers' unions boycotted SATs for year 6 pupils in 2010

---

[14] This is based on an imputed data set prepared by RAND Europe.

- As with all surveys, the figures presented in this report are subject to sampling error. As such, confidence intervals (or margins of error) apply to the quoted statistics (see Appendix D)

- Differences highlighted in the report are statistically significant unless otherwise stated – this does not however necessarily equate to difference having 'real world' significance as large base sizes mean that small differences can sometimes meet the threshold for statistical significance

- The statistical models identify relationships between a range of factors and attainment but we cannot infer causality. A finding might seem to support a particular hypothesis but it is not its absolute proof

Understanding the report and interpreting the findings

This report explores attainment using a series of derived variables which capture common measures such as the pupil's individual circumstances, their experiences, attitudes and behaviour, and the characteristics of schools and areas. In Appendix A we provide some definitions for measures whose derivations are not immediately apparent. Some measures were prepared for the Wave 2 LSYPE2 report on health and wellbeing (RR2) and are described in some depth there; this includes young people's psychological wellbeing measured by the General Household Questionnaire (GHQ12) and two latent constructs; locus of control and 'equates hard work with success'[15].

---

[15] See:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/599871/LSYPE2_w2-research_report.pdf

Chapter 2: Educational attainment and how it varies

The overall level of attainment for any single school cohort receives extensive media coverage when exam results are released. While the accompanying commentary on attainment often looks at broad sub-group differences, for example comparing the performance of boys and girls, the nature of the data available means that highly granular analyses are not possible at the point of publication.

Conversely, the focus of individual pupils and families will typically be much more specific, touching on issues such as whether their own child's results meet their aspirations or constrain their future goals.

The key question we address in this report is quite different; how does attainment vary, on average, between groups of young people with different characteristics and living in different circumstances. To answer this question we consider young people's attainment in terms of two outcome measures. The first of these, our primary outcome measure, is the "Best 8" test score; the young person's capped total points score based on their highest eight GCSE grades including equivalent qualifications. The Best 8 metric uses a continuous scale, technically ranging from zero, perhaps where a young person did not attend their examinations to 464 which equates to eight GCSEs with an A* grade[16].

The distribution of scores in the LSYPE2 cohort (based on weighted data) is shown in Figure 1. The mean score for the full sample, that forms the basis of this chapter, was 319.4 (standard deviation[17]=94.5)[18]. This is slightly higher than the national average of 312.7[19].

---

[16] In practice 0.05% of the sample have scores slightly above 464 which may result from students taking qualifications which are more challenging than GCSEs (e.g. AS levels), or from small data errors.
[17] The standard deviation is a measure used by statisticians and researchers to quantify the amount of variation or spread in scores.
[18] This weighted mean is based on a sample of 9,076 individuals who form the basis of the analysis presented in this chapter.
[19] Data taken from the Statistical First Release (SFR) 01/2016 Table 4a

Figure 1 Distribution of "Best 8" or total capped points based on highest 8 GCSEs or equivalents
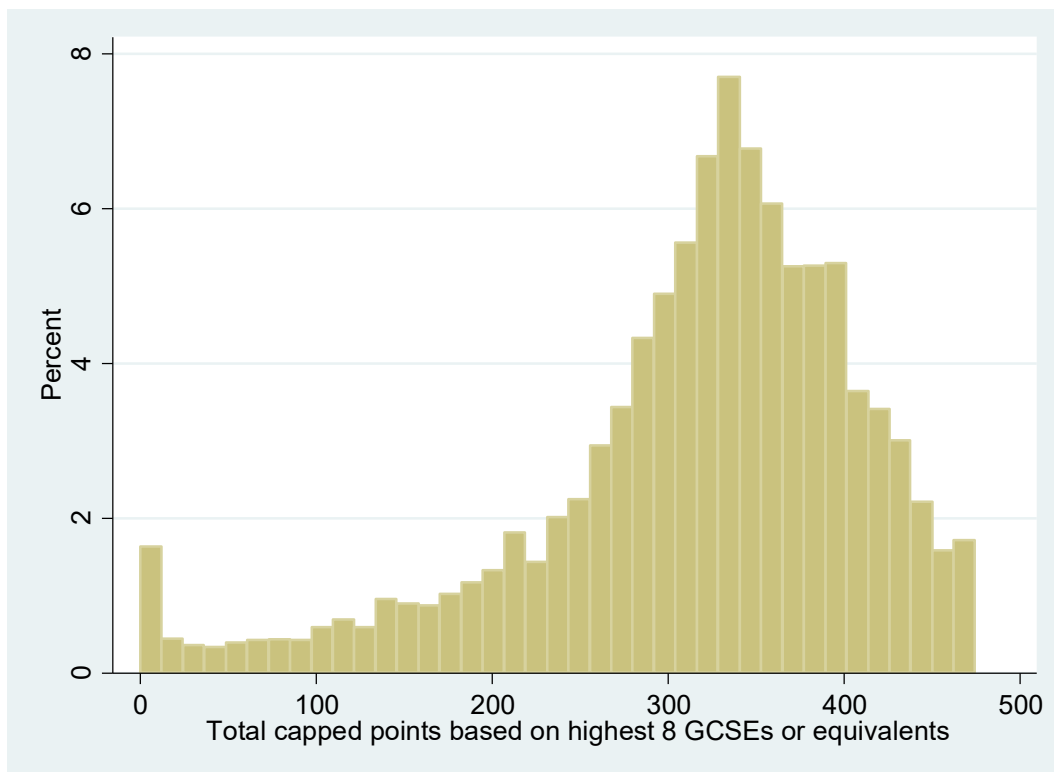
As a general rule of thumb we can treat each 6 points as a grade. So, a rise or fall of 6 points can be seen as moving up or down one grade in one subject and a rise or fall of 18 could be seen, for example, as a 3-grade change in one subject or a 1-grade change in three subjects[20].

---

[20] More specifically, each grade is converted into points on the scale A*=58, A=52, B=46, C=40, D=34, E=28, F=22, G=16. For example, if a student achieved 2Bs, 4Cs, 3Ds and an F, we would include the 2Bs, 4Cs and 2 of the Ds as the best 8 results. The capped points score would be (2*46)+(4*40)+(2*34)=320.

Figure 2 How "Best 8" scores equate for GCSE (or equivalent) grades

If a young person's GCSE attainment was…	Then achieving 6 points fewer (equating to 1 grade) could translate to……			And achieving 24 points fewer (equating to 4 grades) could translate to……	
1 A	1 A	0 As		0 As	1 A
2 Bs	2 Bs	3 Bs		1 B	2 Bs
2 Cs	1 C	2 Cs	or	3 Cs	0 Cs
2 Ds	3 Ds	2 Ds		3 Ds	2 Ds
1 E	1 E	1 E		1 E	3 Es

or between columns 2 and 3: "or"; after column 3: "or a variety of other permutations"; after column 6: "etc."
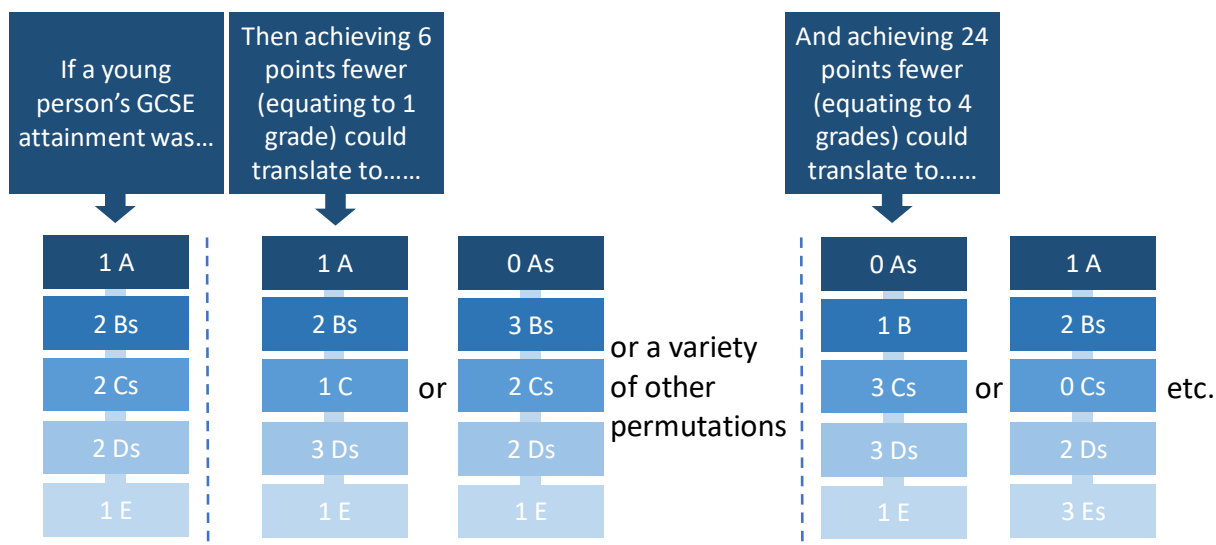
We refer to our second measure of attainment as the "Level 2 English and maths threshold". This level is achieved if the student gains both English and maths GCSEs at grade A*-C (or an equivalent qualification). Overall, the majority of young people (61.7%) were found to have achieved this threshold[21], slightly higher than the national average (59.2%)[22].

---

[21] This is the weighted percentage based on the sample of 9076 individuals that forms the basis of this chapter. In Chapters 3 and 4 we use a sub-sample for young people for whom school information is known and who are not in special schools, and for whom all information is available.
[22] Data taken from SFR 01/2016 Table 4a

Figure 3 Percentage of young people achieving English and maths GCSE or equivalent A*-C
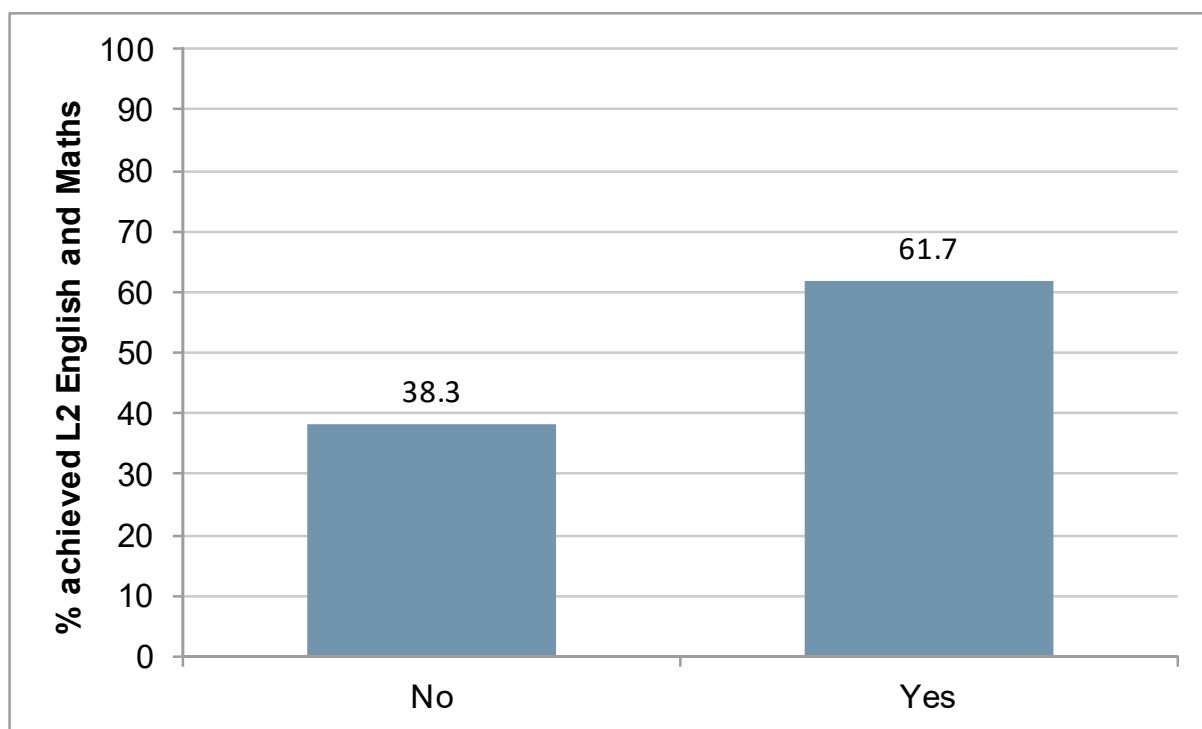
There are, of course, many other ways of measuring attainment, but between them, "Best 8" and the "Level 2 English and maths threshold" provide a good sense of an individual's academic performance. We decided against using the 'A*- C in 5+ GCSEs including English and Maths' measure because it is being phased out. We have also avoided using the most recent measure "Attainment 8[23]" as this wasn't formally introduced until 2016[24].

In this report, we examine how mean Best 8 scores and the achievement of the Level 2 English and maths threshold vary across a wide range of characteristics. We have separated these into a number of domains, as outlined in Figure 4 below. We start by exploring how attainment varies by the personal characteristics and family background of the young people (the dark blue boxes below) and go on to examine how attainment varies by a number of external factors, which relate to the school the young person attends (the turquoise boxes) and the area in which they live (the green box). Later in the chapter we return to this figure and add complexity by introducing some 'intermediary' domains such as the young person's home environment, attitudes and behaviours, health and wellbeing, and their relationship with peers and teachers (which further develops our understanding of school-related factors).

---

[23] The students' average achievement across eight subjects: English, mathematics, 3 other English Baccalaureate (EBacc) subjects (sciences, computer science, geography, history and languages) and 3 further EBacc subjects, GCSEs or approved, high-value arts, academic, or vocational qualification. Source: http://www.theexamsoffice.org/userfiles/files/Progress%208%20and%20Attainment%208.pdf
[24] Details of the most current performance metrics can be found here: https://www.gov.uk/school-performance-tables

Figure 4 Basic domains affecting young people at the individual, school and area level

Area
> **Location and area profile**

School
> **School quality** **School environment**

Individual
> **Personal characteristics** **Family background**
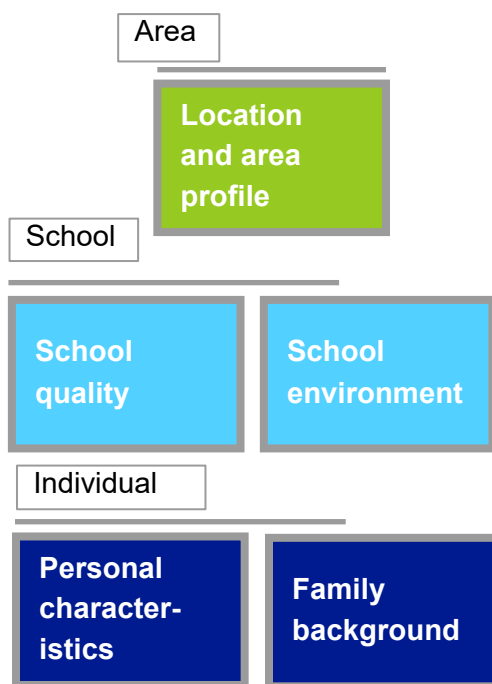
In this chapter, we focus on the main structural characteristics at the individual, school and area level and consider how young people's educational attainment varies by a number of measures within each of these levels in turn. We acknowledge that this initial approach is quite simplistic and does not take account of the overlap between certain characteristics. For example, we show that attainment varies by both ethnic group and by family background (including measures such as household income). However we do not, at this point, look at the interrelationships between these groups, for example by taking account of the considerable variation in the economic position of different Black, Asian and minority ethnic groups which can account for some of their differences in attainment. These more complex considerations are addressed in Chapter 3, where we build statistical models that include all these measures simultaneously, allowing us to understand which are the most important in predicting variations in educational attainment, relative to the others.

As such, **caution is needed when interpreting the much simpler, 'bivariate' analysis, presented in this chapter**. These data show the true variation between different sub-groups in the population, but do not tell the whole story.

Variation by individual characteristics

As has been shown in earlier studies and national statistics[25], outcomes at KS4 vary according to a range of pupil characteristics. In this section we begin by looking at the

---

[25] For example, the 'Revised GCSE and equivalent results in England: 2014 to 2015' SFR published 21 January 2016 highlights very similar patterns to those seen throughout this chapter, underlining the representative nature of the LSYPE2 sample:

areas outlined in the dark blue boxes shown above; the young person's personal characteristics and their family background. Some of these characteristics, like month of birth, are clearly fixed, while others may change as an individual's circumstances shift or their identities form. We treat them here as relatively static so that we can examine, at least superficially at first, how they correspond to differences in young people's attainment.

Personal characteristics

The personal characteristics we consider are: gender; the time of year the young person was born; their ethnicity; whether their parent reports that they have either an illness or disability which affects school; and whether they are identified in NPD records as having a Special Educational Need (SEN).

Gender

Turning first to **gender**, on average girls' attainment was 26.6 points higher than boys' (equivalent to approximately 4 GCSE or equivalent grades). Girls were also more likely to achieve the Level 2 English and maths threshold, with 9.2% fewer boys doing so (67.1% of girls compared with 57.9% of boys). These data are summarised in Figure 5. While these differences are striking, it is important to recognise that the picture is more complex than first appears. Many other factors contribute to the gender differences, as we explain in Box 1 in Chapter 3.

Figure 5 Attainment (total capped GCSE or equivalent points and percentage achieved L2 English and Maths) amongst girls and boys
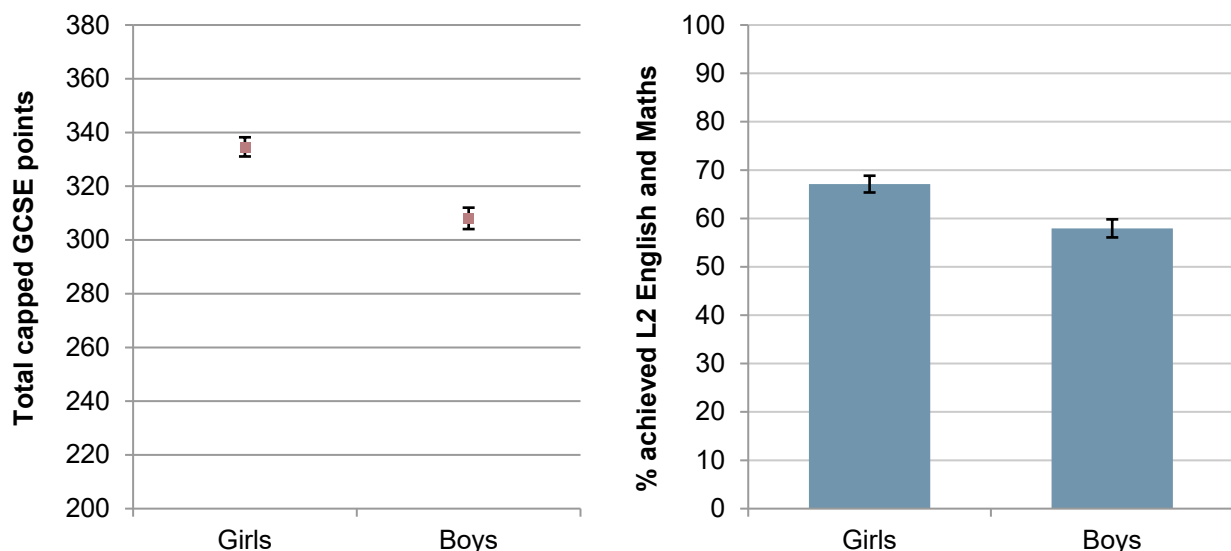
Month and school term of birth

Moving on to **month and school term of birth**, there has been much discussion about the negative impact of being a Summer born baby on attainment[26]. Here we find that young people born between May and August on average achieved lower attainment equivalent to just over one grade (-8.9 points) compared with those born between September and December. Furthermore, 3.6% fewer pupils born in the summer months achieved the Level 2 English and maths threshold than was the case for their counterparts born in the autumn term (60.7% compared to 64.3%).

In this report we focus on school term of birth but, in practice, the pattern is more complex than this. The results here suggest that pupils born in September had the highest level of attainment overall, which was statistically significantly higher on average than children born in all other months except October, March and April, as shown in Figure 6 below[27]. While these differences have caused some concern for parents, a number of studies have improved our understanding of this phenomenon and suggested ways in which it could potentially be addressed, for example by adjusting test-scores but not delaying children's entry to school[28].

---

[26]The 'In-school ability grouping and the month of birth effect: Preliminary evidence from the Millennium Cohort Study' report by Tammy Campbell at IOE provides further discussion of this topic and is also helpful in terms of drawing together existing evidence. See:
http://www.cls.ioe.ac.uk/news.aspx?itemid=2539&itemTitle=Ability+grouping+in+primary+school+may+reinforce+disadvantage+of+summer-born+children%2c+study+finds&sitesectionid=905&sitesectiontitle=Press+Releases

[27] The error bars around each figure show the values between which we are 95% certain the true population value will fall. The potential error in the estimations described by these confidence intervals follows a 'normal distribution', which simply means the true population score is far more likely to be close to estimate than at the edge of the confidence interval. This is the reason why estimates with overlapping confidence intervals may nevertheless be found to have a statistically significant difference from one another

[28] Crawford et al (2013)

Figure 6 Attainment (total capped GCSE or equivalent points) by month of birth
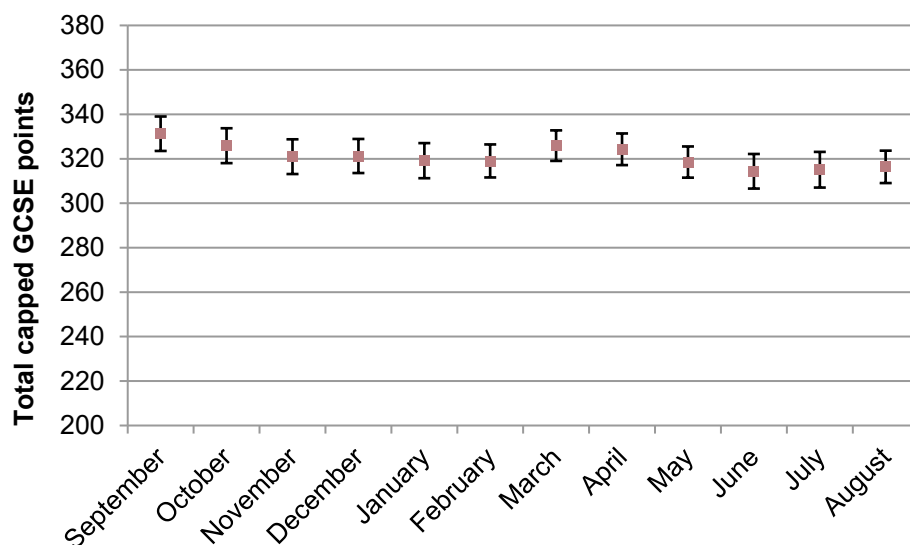
Ethnicity

There was considerable variation in attainment based on pupil's **ethnicity**. Throughout this section we focus on White students as the main comparator group – this is function of their being both a large (and therefore statistically robust) group as well as the fact that White students, on average, sit broadly in the middle of the distribution of attainment by ethnicity. Using the basic approach adopted throughout this chapter, in which scores for each population sub-group are compared directly, we see that on average young people from some Black, Asian and minority ethnic (BAME) groups did significantly better than White students[29]. In particular, Indian students achieved scores almost six grades higher on average (35.1 points), the scores of young people from 'Other' ethnic groups were almost four grades higher (24.6 points) and those from Bangladeshi families were two and a half grades higher (15.6 points) than those of White students on average. That said, other groups, particularly Black Caribbean students, achieved lower average scores (-24.5 points, the equivalent of four grades) than their White counterparts (see Figure 7 below).

Indian students were also far more likely to reach the Level 2 English and maths threshold, with 16.1% more Indian students reaching this threshold than White students (78.4% compared with 62.3%). Two groups were significantly less likely than White

---

[29] The analyses in this report consider White students as a single group. However, other research has identified differences in attainment within the broad 'White' group e.g. Strand, S., Malmberg, L. E., & Hall, J. (2015). *English as an Additional Language (EAL) and educational achievement: An analysis of the National Pupil Database.* London: Educational Endowment Fund. See: http://educationendowmentfoundation.org.uk/uploads/pdf/EAL_and_educational_achievement2.pdf

students to reach the Level 2 English and maths threshold: Black Caribbean (-14.5%); and Pakistani (-6.7%)[30].

These differences are substantial at face value, but can be misleading unless they are considered alongside other factors such as the young person's economic position which varies widely by group and is discussed further in Chapter 3. These data also become more meaningful when we consider other factors that might explain these differences, for example the young person's home environment and their attitudes and aspirations.

In addition to ethnicity we also considered primary language spoken in the home. Initial analysis suggests that there is some variation, with young people who have English as a first language but speak another, or who are classified as bilingual, attaining higher scores at KS4 (10.9 points and 15.5 points higher than English only, respectively - see Figure 7). However, when other factors were taken into consideration in this study, primary language spoken at home did not play a significant role in the variation in attainment and was not included in our main analysis in later chapters.

Figure 7 Attainment (total capped GCSE or equivalent points) by ethnicity and primary language spoken
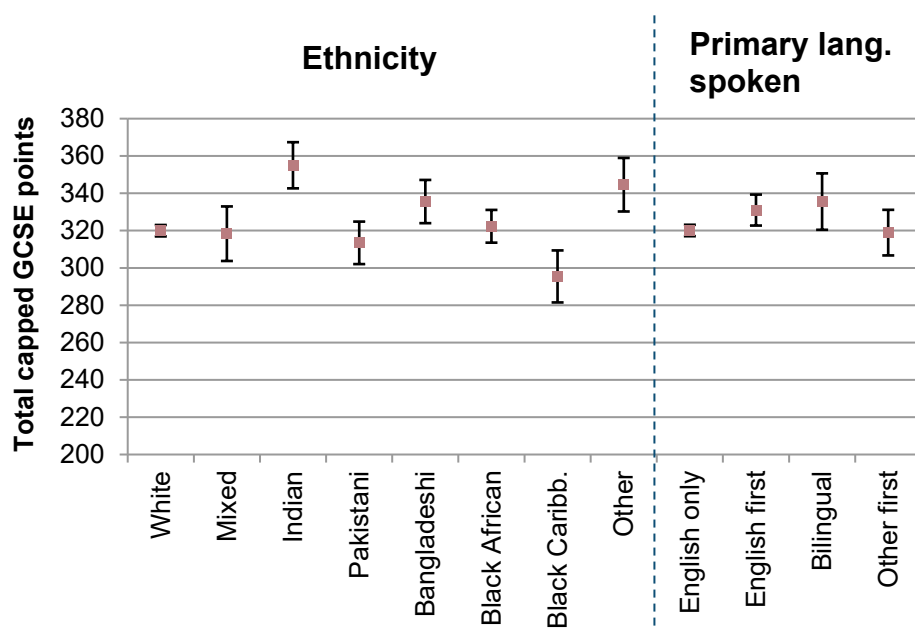
Longstanding illness, disability or infirmity

Young people whose main parent reported that they had **a longstanding illness, disability or infirmity**, whether or not this was reported as affecting their school work, had significantly lower attainment than counterparts whose main parent did not report that they had an illness or disability. However, those with a reported longstanding illness

---

[30] The attainment of minority ethnic groups whose prevalence is low, such as those from a Chinese background and Gypsy Roma Travellers, can be found in official SFR data published by the government alongside whole population data for the groups detailed here.

or disability which affects school experienced the most pronounced disadvantage, scoring on average 115.2 points lower than those with no illness or disability (the equivalent of 19 grades). This is illustrated in Figure 8 below. Of course, the nature of illness or disability contained within this measure will be very broad and the way in which any given illness or disability impacts on an individual may also differ significantly from person to person. Some of this difference may reflect disadvantage and exclusion rather than ability. However, as discussed in Chapter 3, when we later control for other factors, illness or disability which was reported as affecting school corresponded with lower average attainment (while illness or disability which was *not* reported as affecting school did not).

Figure 8 Attainment (total capped GCSE or equivalent points and percentage achieved L2 English and Maths) by disability status
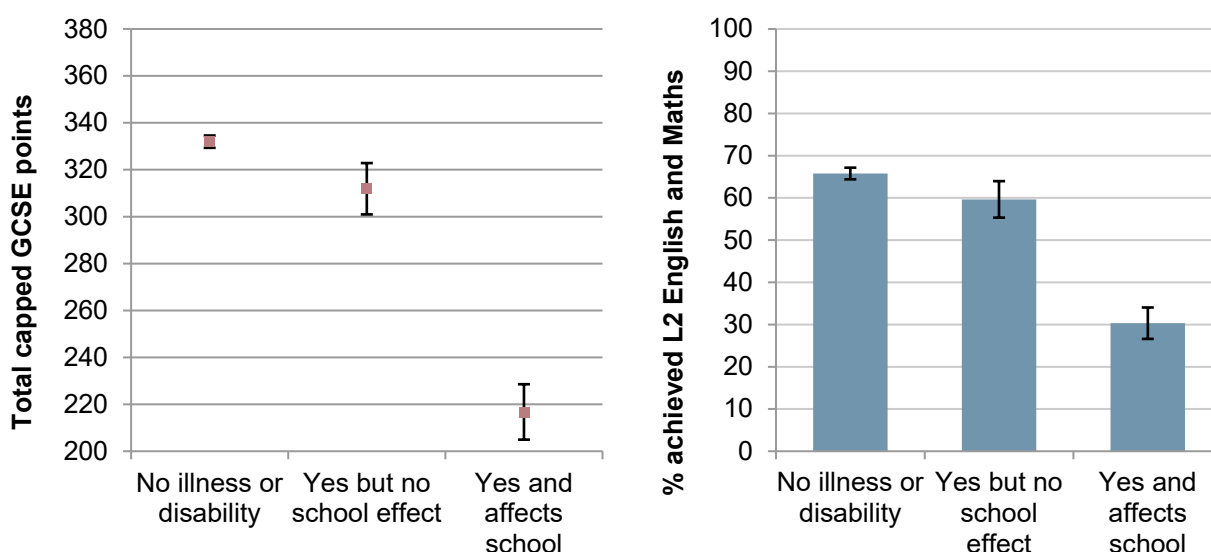
The difference in the percentage of young people in each group achieving the Level 2 English and maths threshold was also very large. There was a 6.1% difference between those with no disability (65.8%) and those reporting a longstanding illness or disability which does not affect their schooling (59.6%). The difference was even more pronounced among for those who reported an illness or disability which affects their schooling, 30.3% of whom achieved this level (35.5% lower than the 65.8% among those who reported no disability). These data are summarised in Figure 8 above (right hand side).

Special Educational Needs

Some young people are identified in the National Pupil Database as having a Special Educational Need (SEN) [31]. Like illness and disability, having a **SEN** (including

---

[31] This means that the young person has a learning difficulty or disability which calls for special educational provision to be made for him or her. A child of compulsory school age has a learning difficulty or disability if he or she has a significantly greater difficulty in learning than the majority of others of the same age, or has a disability which prevents or hinders him or her from making use of educational facilities of a kind generally provided for others of the same age in mainstream schools or mainstream post-16 institutions

statements and/or school action) had a large, negative association with attainment at KS4. Young people with SEN scored 123.5 points lower than the average young person without SEN. Furthermore only 23.6% of young people with SEN achieved Level 2 compared to 71.5% of those without SEN, a difference of 47.9%[32]. It was to be expected that attainment for this group would be lower, given the challenges which young people with SEN must overcome. These data are shown in Figure 9.

Figure 9 Attainment (total capped GCSE or equivalent points and percentage achieved L2 English and Maths) by SEN status (including statements and/or school action)
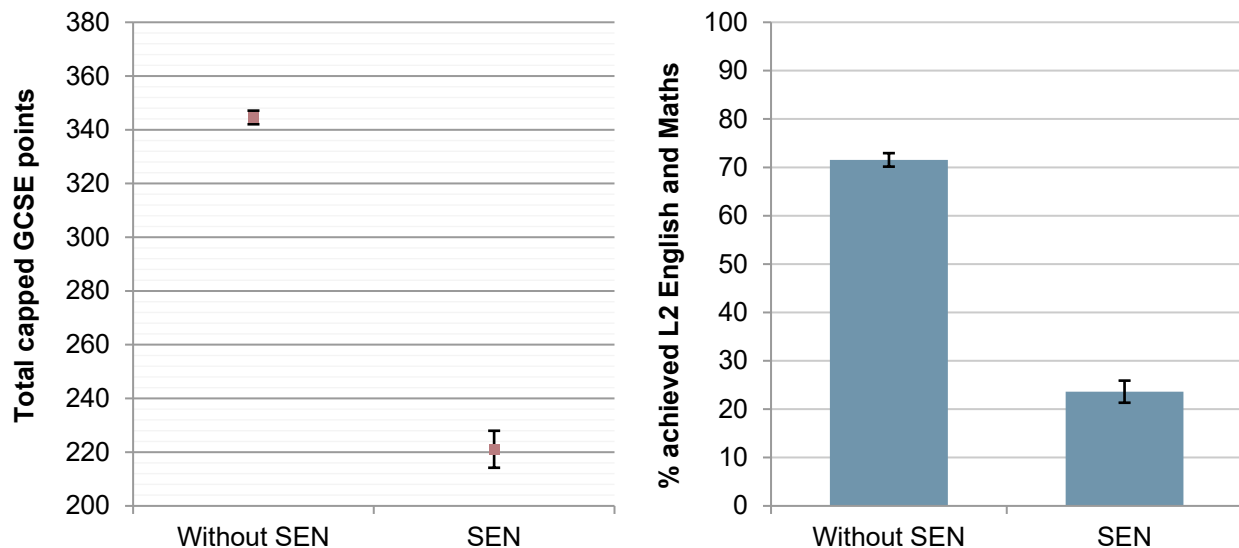
Family background

As shown above, educational outcomes vary, on average, with personal characteristics. However, attainment is also known to vary significantly according to the young person's family background and their social and economic circumstances[33]. For this report we consider a number of such factors, including: the composition of their family; whether any parent-figures in the household were working (in full or part time paid employment or self-employment); and the highest qualification held by the young person's mother. Family background can also be thought about directly or indirectly in terms of markers of social economic position, for example with regard to their socio-economic group (or NS-SEC), their household tenure and their household financial situation. We have summarised household finances broadly in terms of whether they were currently eligible for free school meals, had been eligible in the last six years and, for those who were above this

---

[32] The figure based on the sample is 71.5% but is 66.4% based on the national figure as reported in SFR.
[33]See:
Strand, S. (2011). The limits of social class in explaining ethnic gaps in educational attainment. *British Educational Research Journal, 37*(2), 197-229. doi:doi:10.1080/01411920903540664
Strand, S. (2014). Ethnicity, gender, social class and achievement gaps at age 16: intersectionality and 'getting it' for the white working class. *Research Papers in Education, 29*(2), 131-171. doi:10.1080/02671522.2013.767370

threshold, whether their total household income was above or below the UK median household income[34].

Family composition

There were significant differences in attainment depending on **family composition**. Living with both biological parents was associated with higher scores at KS4 than other types of family composition. Young people living with one biological parent (either in a single parent household or with an additional step parent) achieved lower average scores (-39.6 and -42.5 points respectively) than those living with both parents. Young people who were not living with either biological parent[35] had substantially lower attainment (-105.7 points) than those living with both parents. These data are shown in Figure 10. This finding is consistent with the large deficit for looked after children that is reported in administrative data (DfE 2016a, DfE 2016b).

Young people living with one biological parent in a single parent family were also much less likely to achieve the Level 2 English and maths threshold (53.4% compared with 68.3% amongst those with two biological parents), as were those living in a step family (51.3%). The difference was even more pronounced for young people in households with no biological parents, 35.1% of whom achieved the Level 2 English and maths threshold (33.3% fewer than was the case among those living with both biological parents) [36].

---

[34] Which in 2014/15 was a net income £25,600. Further details of this measure are discussed later in this section.

[35] These young people live with a range of different people including adoptive and foster parents and other family members.

[36] Nationally, only 16% of young people who are 'looked after' achieved this threshold. Most of these are foster placements, but this group also includes young people in secure units, children's homes, hostels and so on, but not informal family arrangements. The figures based on the survey cannot be directly compared but it is clear that our sample of young people with no biological parents are more academically successful than the national sample of looked after children.

Figure 10 Attainment (total capped GCSE or equivalent points) by family composition
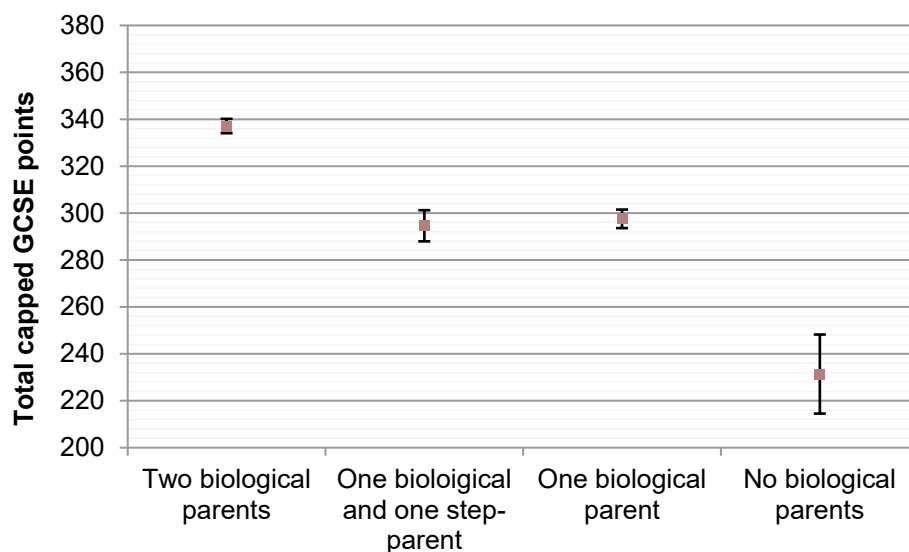
Parental employment status

Turning to **parental employment status**, in households with full employment[37] young people did significantly better at KS4 than their counterparts. In two parent households, in which only one parent was employed, young people achieved 25.4 points less on average than those in households where all parents were employed (equivalent to approximately 4 grades). Young people in households in which no parents were employed achieved 79.2 points less on average than those in households with full employment (equivalent to approximately 13 grades). A similar picture is true for the proportion of young people achieving the Level 2 English and maths threshold. Young people in two parent families where only one parent worked were less likely to meet this threshold than those living in fully employed households (57.1% compared with 67.7%, a difference of 10.6%). The proportion of young people meeting this threshold in households in which no parent was working was markedly lower (37.6%, a difference of 30.1% compared to those in households with full employment).

Past studies have shown that parental education – and particularly maternal education – is an important predictor of pupil attainment[38]. The present study also demonstrates a very strong gradient in pupil attainment according to the qualification level of the mother. Young people whose mother had achieved a further education qualification below degree level achieved a score 30.1 points lower than those whose mother held a degree (equivalent to approximately 5 grades). Where a mother's highest qualification was at A Level young people achieved 44.8 less points on average (equivalent to approximately 7 grades) than those whose mother held a degree. As shown in Figure 11 below, the

---

[37] Defined as two parent households where both parents were in either full or part time paid employment or self-employment OR one parent households where the parent was in either full or part time paid employment.
[38] For example: **2017/14** Sullivan, A, Moulton, V, Fitzsimons, E (2017) 'The intergenerational transmission of vocabulary' CLS working paper 2017/14. London: Centre for Longitudinal Studies

attainment of young people continued to decline with each subsequent reduction in the level of qualifications held by the mother. It should again be noted that these data do not take into account the interrelationships between variables which are discussed in Chapter 3 of this report.

Figure 11 Attainment (total capped GCSE or equivalent points) by maternal qualifications
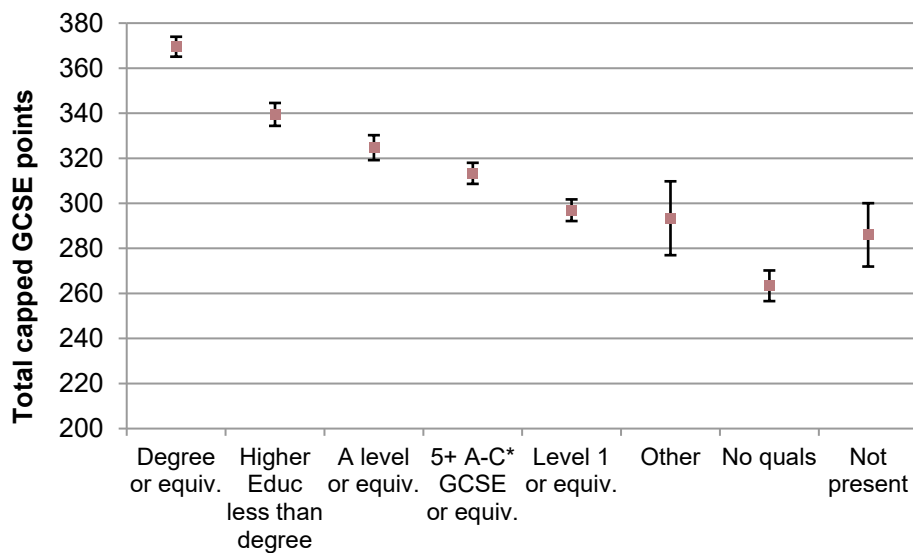
In terms of achievement of the Level 2 English and maths threshold, a similar pattern emerges - 81.4% of young people whose mother's held a degree achieved this threshold compared with only 37.3% of young people whose mothers held no qualifications (see Figure 12).

Figure 12 Attainment (percentage achieved L2 English and Maths) by maternal qualifications
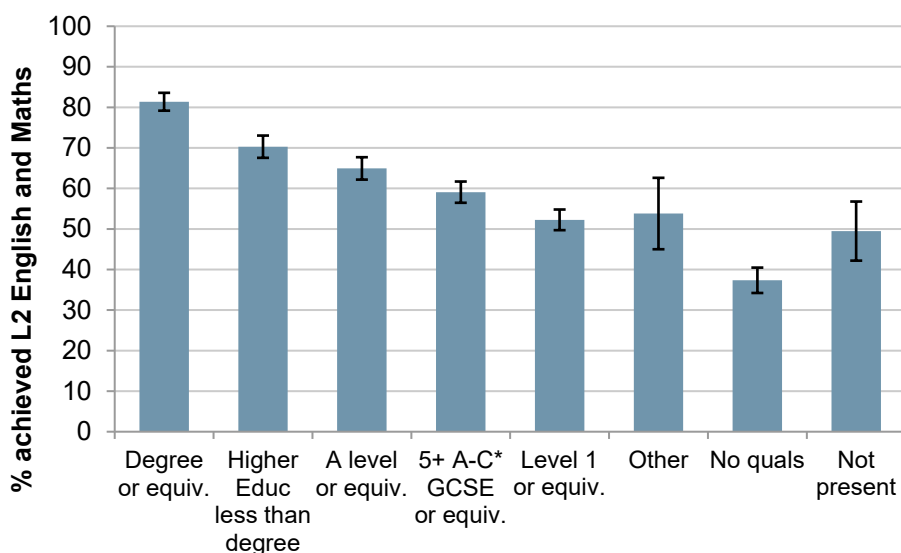
Socio-economic Classification

The **National Statistics Socio-economic Classification (NS-SEC)** is the official socio-economic classification in the United Kingdom. Broadly speaking, it is a measure of household employment and occupational status[39]. We have based the measure of NS-SEC used in this report on the highest reported NS-SEC status in the household, which could be either the mother's or father's (or male or female guardian's) status. Similar to maternal qualifications, there was a clear attainment gradient related to NS-SEC, as shown in Figure 13. Attainment was highest for young people living in managerial or professional occupation households, and reduced across the NS-SEC categories: young people from intermediate occupation households achieved 36.1 less points on average (equivalent to approximately 6 grades less); those from routine and manual occupation households achieved 77.3 less points on average (approximately 13 grades less); and those from households where no parent was working achieved the lowest scores (82.8 less points on average, equivalent to approximately 14 grades)[40].

A similar gradient was evident when looking at the achievement of the Level 2 English and maths threshold. Overall, 75.7% of young people in managerial or professional occupation households achieved this threshold. The equivalent proportion of young people in intermediate households was 15.8% lower (59.9%); in routine and manual occupation households it was 32.8% lower (42.9%); and in households where no parent was working it was 37.7% lower (38.0%).

---

[39] More details about the derivation of NS-SEC can be found here: http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/soc2010-volume-3-ns-sec--rebased-on-soc2010--user-manual/index.html

[40] Appendix D contains tables showing how educational attainment varies (for both Best 8 and the percentage of young people achieving the Level 2 English and maths threshold) when using the more common, eight-category version of NS-SEC. The observed pattern is very similar but we present the 4 category variable here because it is the version used in the models in Chapters 3 and 4, which had a clearer association with the outcome.

Figure 13 Attainment (total capped GCSE or equivalent points and percentage achieved L2 English and Maths) by the National Statistics Socio-economic Classification
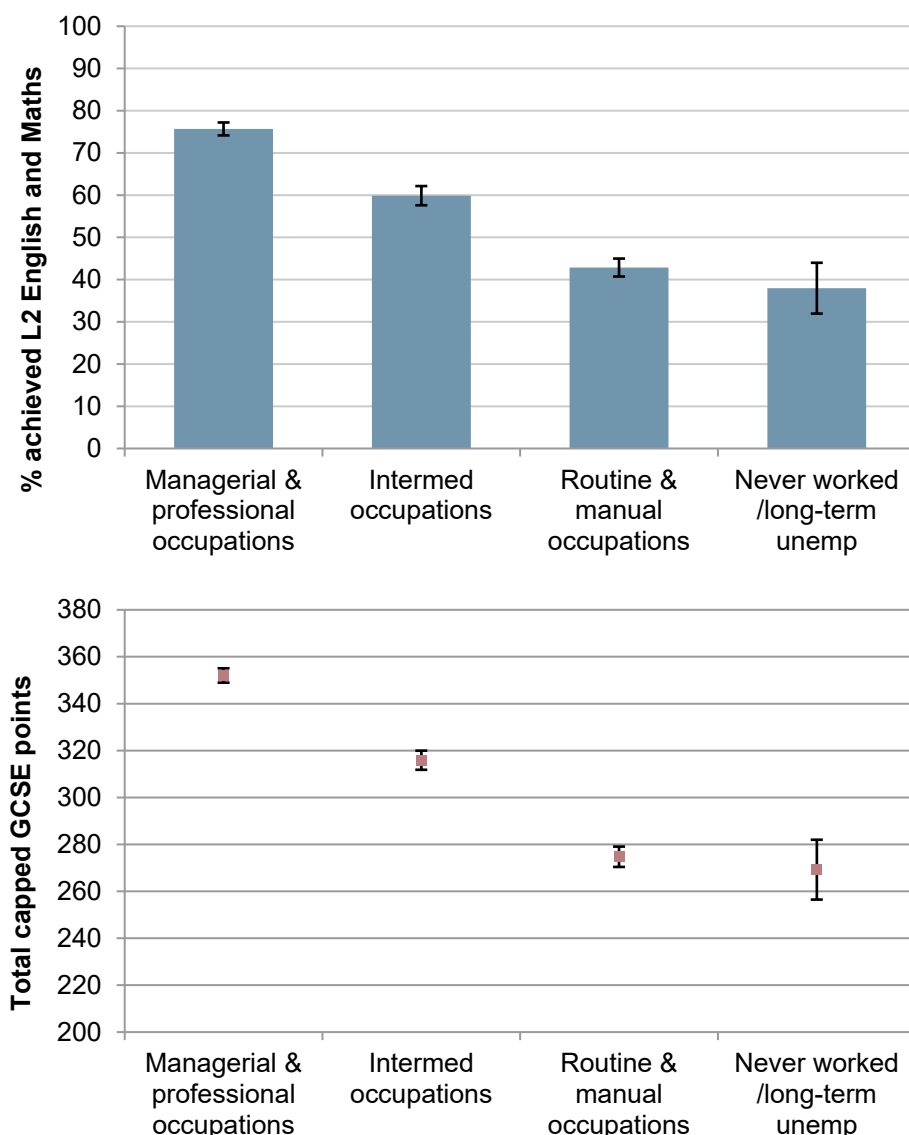
Eligibility for Free School Meals and household income

Turning now **to FSM and indicators of household income**, a derived measure was created combining household income and FSM status, as has been done in earlier studies (Strand, 2014b). This combines an estimate of whether the young person lives in a household above or below the median UK household income[41], and whether the young person is currently eligible for FSM, or has been in the last six years. Such an approach

---

[41] Which in 2014/15 represented a net income £25,600. For further details please see: https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/bulletins/nowcastinghouseholdincomeintheuk/2015-10-28. For the derivation of this measure, an indicative proxy household income measure has been used, as LSYPE2 records only gross household income within income bands. The closest income band to the UK net median is '£26,000 - £31,199 gross household income'. If a household earns, this income or below they have been classified as 'Below the median household income', if a household earns more than this income, they have been classified as 'Above the median household income'. Combining these two variables reduces the amount of missing data and provides some differentiation among the non-FSM population.

reduces the number of cases with missing data and also provides some differentiation among the non-FSM population. Four categories were created: (1) Currently eligible for FSM; (2) Not currently eligible for FSM, but has been eligible for FSM in the last six years; (3) Having an estimated household income less than the UK median and not eligible for FSM in the last six years; (4) Having an estimated household income more than the UK median and not eligible for FSM in the last six years.

In common with both mother's education and household NS-SEC, there was a strong attainment gradient associated with economic circumstance. Young people living in households with above median income (and who had not been eligible for FSM in the last six years) achieved an average 355.5 points at KS4. When we compare this group with young people from more deprived backgrounds we find that:

- those from households with a lower than median yearly income, that had not been eligible for FSM in the last six years, scored 43.6 fewer points on average (approximately 7 grades);
- those from households where the young person had been eligible for FSM in the last six years, but were not currently eligible, scored 66.3 fewer points on average (approximately 11 grades); and
- those from households where the young person was currently eligible for FSM scored 94.1 fewer points on average (approximately 16 grades);

These findings are illustrated in Figure 14.

Figure 14 Attainment (total capped GCSE or equivalent points and percentage achieved L2 English and Maths) by FSM and household income
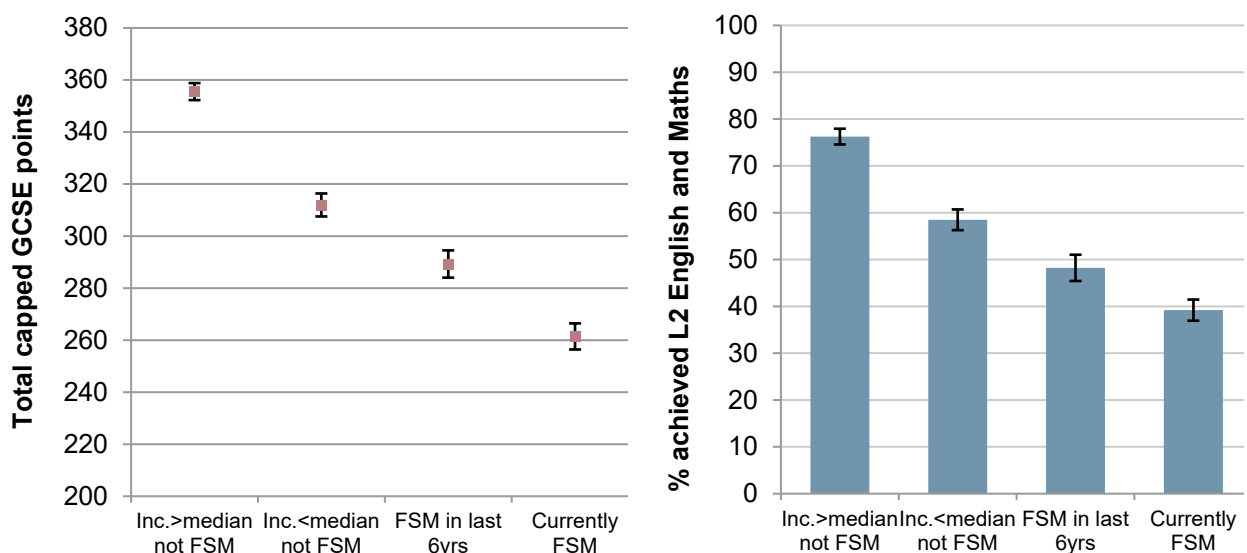
The percentage of young people achieving the Level 2 English and maths threshold also declined substantially across these same categories, from 76.3% of young people in

households with an estimated household income above the UK median to 39.2% of young people who were currently eligible for FSM[42].

Housing tenure

Housing tenure is also often used as an indicator of the socio-economic circumstances of a household, but can also reflect housing security among other things. Here we find that young people who lived in a home which was owned outright or with a mortgage achieved the highest attainment on average (343.4 points). As shown in Figure 15, attainment was lower amongst those living in council and housing authority rental properties (by -76.5 points, approximately 13 grades), those in private rental properties (by -47.6 points, approximately 8 grades) and those in 'other' accommodation types (by -27.7 points, approximately 5 grades) [43]. Tenure was also associated with the achievement of the Level 2 English and maths threshold. This standard was achieved by 72.2% of young people in houses owned outright or on a mortgage, decreasing to 38.6% of those in properties rented from the council or housing association, 51.6% of those in private rental properties and 56.2% of those in 'other' accommodation types.

Figure 15 Attainment (total capped GCSE or equivalent points) by housing tenure
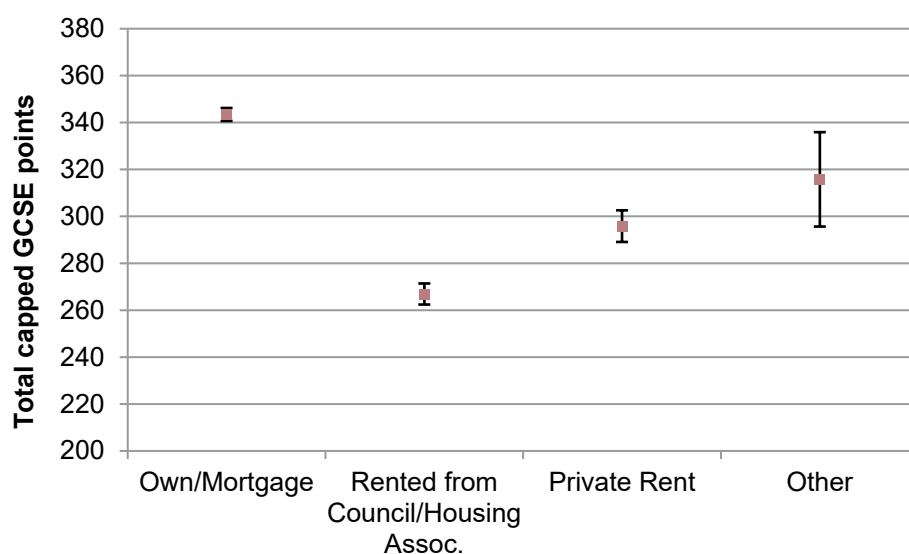
Variation by school characteristics

So far we have considered young people's personal characteristics and their family background (the two dark blue boxes in Figure 4 on page 27). As discussed earlier, there

---

[42] The proportion in the population for FSM pupils was 35.7% suggesting that the students in the analytical sample had slightly higher attainment than the national equivalents.

[43] 'Other' accommodation types include a range of less typical housing circumstances including, amongst others, those living with family or friends and those living in other rent-free circumstances. Those living in hotels, B&Bs and institutions were not included in this analysis.

are other structural factors which are likely to impact on a young person's attainment. Here we look at how attainment varies by school level characteristics (the turquoise boxes in Figure 4). In particular we examine the structural characteristics and composition of the school along with measures of school quality.

Structural characteristics and composition of the school

A broad range of measures which describe the structural characteristics and composition of the school are presented in this section: whether the school has a sixth form; whether it is single-sex or co-educational; whether it is selective; its size; its composition in terms of ethnicity; the prevalence of students eligible for free school meals; and the prevalence of students who speak English as a second language. We do not consider school type (for example, whether the school is an Academy), since the status of individual schools was changing rapidly at the time of the survey and findings would therefore be difficult to interpret.

School intake

Firstly exploring whether the young person attends a **school with a sixth form**, whether their school is **single or mixed** gender, and whether they attend **a selective school,** we find the following:

- **Going to a school with a sixth form** was related to attainment. As shown in Figure 16, young people who attended schools without a sixth form achieved lower scores on average (13.3 points lower, the equivalent of approximately 2 grades). They were also less likely to reach the Level 2 English and maths threshold (by 4.6%, with 58.9% reaching the threshold compared with 63.5% in schools with sixth forms). These differences are relatively small but are nevertheless statistically significant.

- **Gender composition of school** - there was a relationship between the gender composition of the school and attainment. On average, young people who attended mixed schools achieved lower scores than those attending single sex schools (36.6 points lower on average, approximately 6 grades less). Those in mixed schools were also less likely than those in single sex schools to have achieved the Level 2 English and maths threshold (an 11.7% difference). Once again, it should be noted that the findings in this chapter look at each variable in isolation – in the later chapters, where the data is modelled to take account of multiple factors at the same time, gender composition was not found to be strongly associated with attainment.

- As might be expected, attending **a selective school** was strongly associated with attainment. On average, young people who attended non-selective schools achieved lower scores (on average 89.7 points lower, a difference of approximately

15 grades). There was also a pronounced difference in terms of the proportion of young people achieving the Level 2 English and maths threshold. In non-selective schools, 61.6% of young people reached the threshold, compared with 97.6% of students in selective schools (a difference of 36.1%). When considering attainment at selective schools it is important to bear in mind not only the relatively positive starting point of pupils attending such schools, but also the collective impact of selection on the whole cohort of young people. For example, the Education Endowment Fund draws attention to the negative impacts of setting or streaming on lower achieving students[44].

Figure 16 Attainment (total capped GCSE or equivalent points) by whether the school has a sixth form, is single or mixed sex and is selective or non-selective[45]
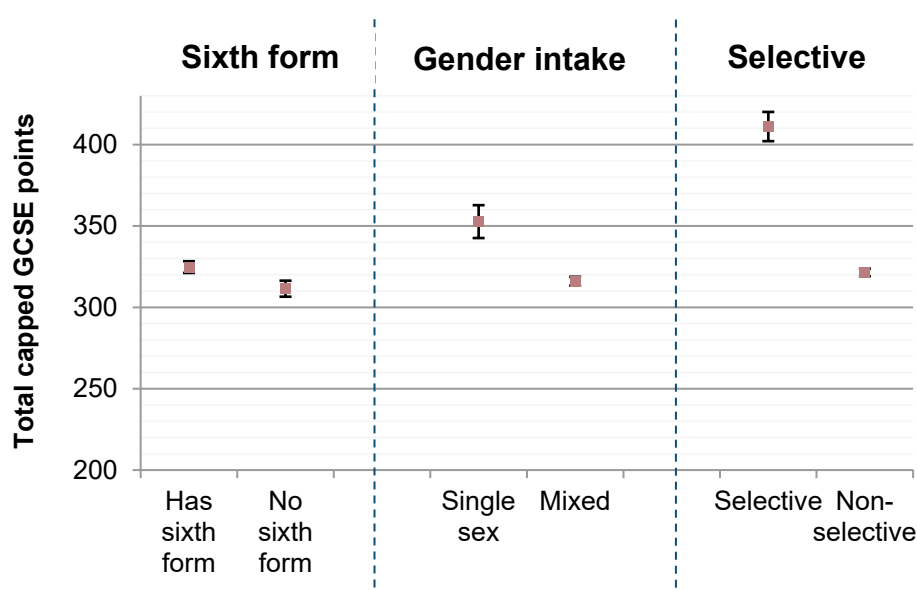
School Size and composition

For simplicity, when describing school size and composition we use quintiles (for example, breaking down schools into five categories running from the largest to the smallest) [46].

We begin with **school size** (defined according to the number of full time equivalent pupils). As Figure 17 shows, on average pupils attending the smallest schools had lower levels of attainment than those attending larger schools[47]. For example, pupils at schools

---

[44] https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/setting-or-streaming/

[45] Please note the y axis scale on this chart is different to those used elsewhere in this chapter (with a view to accommodating an atypically high value).

[46] These quintiles are based on the LSYPE sample rather than the whole school population.

[47] For the purpose of describing this relationship, school size has been divided into quintiles based on the LSYPE sample with the following number of pupils: Quintile 1=24-771 pupils; Quintile 2=772-983 pupils; Quintile 3=984-1194 pupils; Quintile 4=1195-1448 pupils; and Quintile 5=1449-2640 pupils.

in the smallest quintile achieved scores that were 43.2 points lower on average (approximately 7 grades) than those in schools in the next quintile. A similar pattern is seen when examining the effect of school size on achievement of the Level 2 English and maths threshold, with young people in the smallest quintile of schools being less likely to achieve this threshold.

Figure 17 Attainment (total capped GCSE or equivalent points) by number of pupils in school
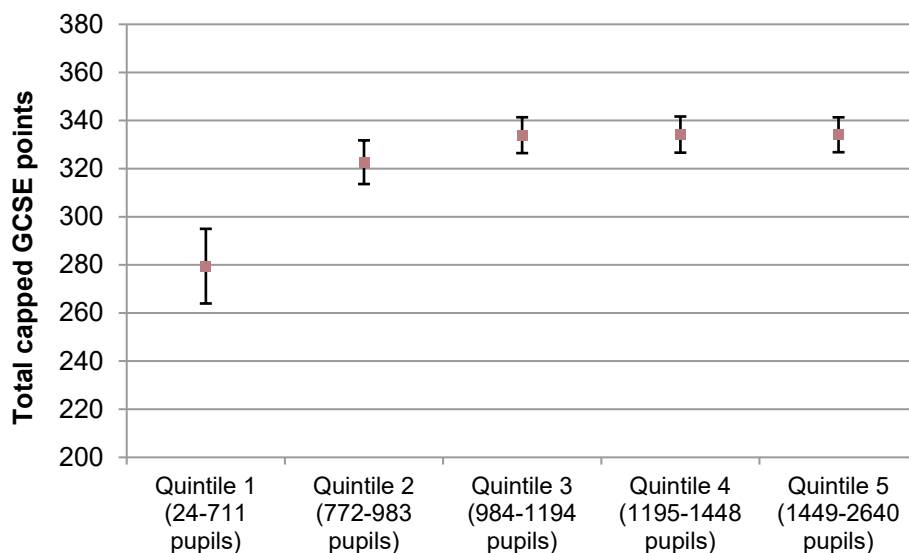
There was also a significant relationship between the proportion of SEN pupils (including statements and/or school action) and the attainment of pupils attending the school[48]. However, it is important to bear in mind that in this chapter we are looking at factors in isolation. When we look at the data more holistically and other characteristics are also taken into account in later chapters, the relationship between the proportion of SEN pupils and attainment was not found to be statistically significant.

Young people who attended schools in the highest two quintiles (with 2% or more SEN pupils) were more likely to have lower attainment scores than those attending schools in the lowest quintile. For example, pupils at schools in the highest quintile (which includes a small number of schools with very a high proportion of SEN students) achieved scores that were 45.7 points (more than 7 grades) lower on average than those in schools in the quintile with the fewest SEN students (below 0.9%). A similar pattern was evident when examining the proportion of pupils who achieved the Level 2 English and maths threshold.

---

[48] Again for the purpose of describing the relationship, the proportion of SEN pupils has been divided into quintiles based on the LSYPE sample, with the following number of pupils: Quintile 1=0-0.8%; Quintile 2=0.9-1.3%; Quintile 3=1.4-1.9%; Quintile 4=2-2.9%; and Quintile 5=3-100%.

There was a particularly strong relationship between the percentage of FSM eligible pupils attending the school and attainment[49]. It is again important to bear in mind the fact that the data in this chapter look at each variable in isolation whereas the data in later chapters adopt a more holistic approach and other characteristics are also taken into account.

Young people attending schools with a higher proportion of pupils eligible for FSM achieved lower average scores and were less likely to have achieved the Level 2 English and maths threshold. Pupils who attended schools in the lowest quintile (with the fewest pupils eligible for FSM) achieved on average 360.2 points at KS4, while those in the highest quintile achieved an average of 269.2 points (a difference of 91 points, the equivalent of around 15 GCSE grades). These data are summarised in Figure 18

Figure 18 Attainment (total capped GCSE or equivalent points) by proportion of pupils eligible for FSM in school
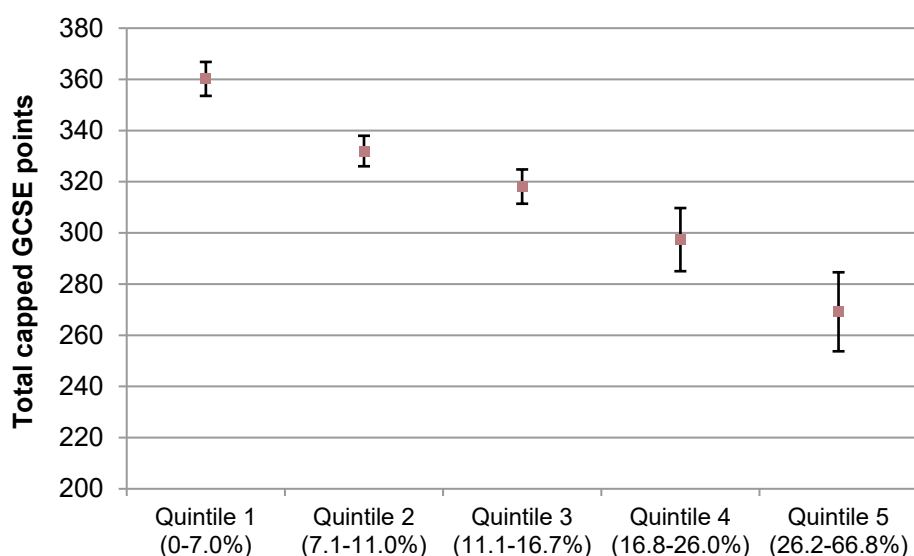
The last two variables relating to school composition were the **percentage of pupils with English as an additional language (EAL)** and the **ethnic composition** of the pupils attending the schools (measured as the proportion of non-white pupils attending the school). The percentage of EAL pupils in a school had no bearing on pupil attainment, and its ethnic composition only a very small effect.

School quality

In addition to the factors relating to the structural characteristics and composition of schools (discussed above), a range of measures which indicate school quality have been

---

[49] For this analysis, the proportion of pupils who attend the school who are eligible for FSM has been divided into quintiles based on the LSYPE sample with the following proportions: Quintile 1=0-7.0%; Quintile 2=7.1-11.0%; Quintile 3=11.1-16.7%; Quintile 4=16.8-26.0%; and Quintile 5=26.2-66.8%.

examined: the ratio of pupils to teachers; a measure of 'school added value'; the level of unauthorised absence in schools; and school OFSTED rating[50].

There is a relationship between **pupil-teacher ratio** and attainment. Young people attending schools in the quintile with the lowest pupil to teacher ratio (1 teacher to 13 pupils or less) attained 50.3 fewer points on average than those in the next quintile. A similar finding is evident with achievement of the Level 2 English and maths threshold, where 47.4% of young people in schools in the quintile with the lowest pupil-teacher ratio achieved this threshold compared to 63.5% in the next quintile, a difference of 16.1 percentage points. It is likely that most of this effect is driven by a small number of schools with very low teacher to pupil ratios which reflect the special needs of their pupils.

There was also a strong relationship between the **OFSTED rating** of the school attended (all of which were maintained schools and therefore had OFSTED ratings) and attainment. Young people attending schools with a rating of ''outstanding' scored an average of 32.6 more points (equivalent to around 5 grades) more than pupils attending schools rated 'good'. This better performance of young people in schools rated outstanding was even more pronounced when compared to schools with a rating of 'requires improvement' (45.6 points, 8 grades) and 'inadequate' (59.3 points, 10 grades), as shown in Figure 19. The same was true for the proportion reaching the Level 2 English and maths threshold (with 74.9% of those from 'outstanding' schools reaching this threshold, compared to 62.6% of those from 'good' schools (a difference of 12.3%), 53.6% of those from schools rated as 'requires improvement' (a difference of 21.3%) and 50.3% of those from 'inadequate' schools (a difference of 24.6%). It should be noted that the relationship between OFSTED rating and attainment is circular to some extent, given that attainment constitutes a part of the OFSTED assessment criteria[51].

When examining **the level of unauthorised absence in schools**, we found that attainment at KS4 was substantially higher in schools with lower levels of unauthorised

---

[50] This report uses the 'overall effectiveness' Ofsted rating, which has the following categories: (1) Outstanding; (2) Good; (3) Requires improvement; (4) Inadequate. In making their judgements about a school's overall effectiveness, inspectors will consider whether the standard of education is good or whether it exceeds good and is outstanding. If it is not good, then inspectors will consider whether it requires improvement or is inadequate. In judging the overall effectiveness, inspectors will take account of the four key judgements of the following: effectiveness of leadership and management; quality of teaching, learning and assessment; personal development, behaviour and welfare; and outcomes for children and learners. Before making the final judgement on the overall effectiveness, inspectors must evaluate: the effectiveness and impact of the provision for pupils' spiritual, moral, social and cultural development; and the extent to which the education provided by the school meets the needs of the range of pupils at the school including pupils who have disabilities and pupils who have special educational needs. Further details can be found here:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/654625/School_inspection_handbook_section_5.pdf
[51] Current inspection guidance can be found at https://www.gov.uk/government/publications/school-inspection-handbook-from-september-2015

absence. As shown in Figure 19, a similarly strong gradient to that associated with Ofsted ratings was evident. On average, young people in schools in the quintile with the worst unauthorised absence record[52] scored 56.4 points less (approximately 9 grades) than those in schools in the top quintile, with up to 0.6% of school sessions missed through unauthorised absence. The same pattern was true for proportion of young people reaching the Level 2 English and maths threshold with only 50.4% of young people in schools with the worst unauthorised absence record reaching the threshold compared to 72.4% in schools with the best, a difference of 22%.

Figure 19 Attainment (total capped GCSE or equivalent points) by School Ofsted Rating and unauthorised absence rating
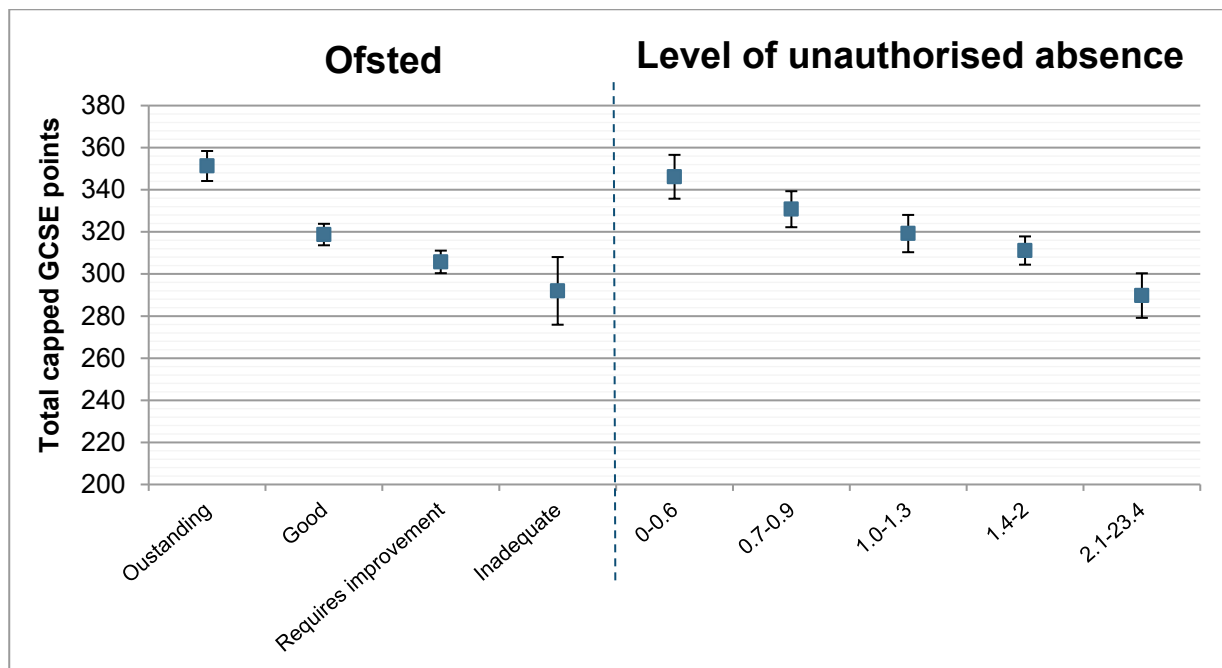
Variation by area level characteristics

We now carry out a similar analysis of attainment examining area level characteristics. These are individual level characteristics based on the location where the young person lives[53]. We firstly explore locality (based on region and whether they lived in an urban or rural area), and then move on to examine areas with differing levels of deprivation.

---

[52] This quintile contains a very wide range of schools, starting with those with 2.1% of school sessions missed through unauthorised absences but also including small numbers of schools with up to 23.4% of sessions missed in this way.

[53] This set of variables provides more relevant and accurate information about the young persons' locality than considering Local Education Authority. We considered the alternative of using school postcodes but this would not have provided varied information for young people attending the same school but living in different circumstances.

Locality

The **regions of England** (formerly known as Government Office Regions) are the highest tier of sub-national division in England and are used as region-level geography for statistical purposes. There are nine regions across England. There were significant differences in average pupil attainment at KS4 across some regions of England, based on the location where that young person lived. In comparison to young people living in London, attainment of young people is significantly lower in the following regions: North East (-14.4 points); East Midlands (-14.8 points); North West (-16 points); West Midlands (-18.3 points); and Yorkshire and the Humberside (-24.3 points). This is illustrated in Figure 20. The same five regions showed the largest attainment deficits compared to London in the official published data on GCSE attainment[54]. Again it should be noted that later chapters look at the data more holistically to understand the factors underlying such differences.

Figure 20 Attainment (total capped GCSE or equivalent points) by Region
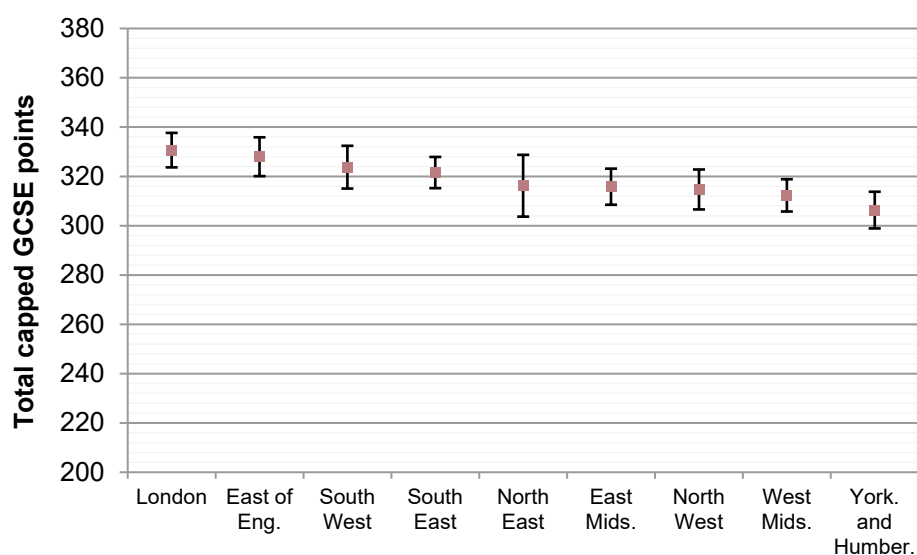
Turning to the achievement of the Level 2 English and maths threshold, whilst some differences were evident, these were less widespread within the LSYPE2 sample. In comparison to London, where 64% of young people achieved the threshold level, young people from the following three regions were found to be less likely to achieve this standard:

- those from the North West (58.3% achieved the threshold, a difference of 5.7%),
- those from the West Midlands (58.5% achieved the threshold, a difference of 5.5%),

---

[54] The official capped KS4 scores for each region in 2015 were: London (326.4), South East (317.9), South West (315.6), East (314.7), North West (310.3), West Midlands (308.8), North East (307.4), Yorkshire and Humber (305.7) and East Midlands (305.2). See: https://www.gov.uk/government/statistics/revised-gcse-and-equivalent-results-in-england-2014-to-2015

- those from Yorkshire and Humberside (56.7% achieved the threshold; a difference of 7.3%)[55].

We also looked at attainment differences between young people living in **urban and rural areas**. However, there were no significant variations by this measure in terms of attainment at KS4 (318.0 points and 324.7 points respectively), or in terms of the proportion of young people achieving the Level 2 English and maths threshold (61% and 64.1% respectively).

Area type

Following on from this analysis it is also important to consider the characteristics of the areas in which each young person lives, because it may not be region per se that is impacting on attainment, but rather differences in the characteristics of the regions. Such effects will be more fully taken into account in the multi-variate analysis in Chapter 3, but it is perhaps valuable to start by illustrating how the differing levels of deprivation found across the country might come into play.

There are a number of different ways of characterising area based on different dimensions captured within the Index of Multiple Deprivation (IMD). We considered a variety of approaches to representing this within the study but ultimately decided to use the **Income Deprivation Affecting Children Index (IDACI)** here and throughout the report as a measure of deprivation[56,57]. Our analysis was based on IDACI quintiles and found that there was a strong relationship between IDACI and attainment. Attainment at KS4 was highest amongst young people in the top IDACI quintile and performance then declined across the IDACI quintiles, with young people in the lowest IDACI quintile (those living in the 20% most deprived areas) achieving -74.7 points lower on average (approximately 12 grades less) than those in the highest IDACI quintile (those living in the 20% least deprived areas). Figure 21 illustrates this pattern. The same pattern is evident when examining achievement of the Level 2 English and maths threshold. In the

---

[55] While the differences between London and the North East / East Midlands were not statistically significant in the LSYPE2 data, administrative sources suggest that these regions also show a performance deficit relative to London in terms of the Level 2 English and maths threshold. See Table LA1: https://www.gov.uk/government/statistics/revised-gcse-and-equivalent-results-in-england-2014-to-2015

[56] IDACI is a measure of deprivation which is based on the proportion of children aged under 16 living in low income households in different areas of the country. Young people who fall into the first IDACI quintile are those who live in the 20% of areas with the lowest proportion of children in low-income households (i.e. they live in the least deprived areas). Those in the fifth IDACI quintile live in the 20% of areas with the highest proportion of children in low-income households (i.e. they live in the most deprived areas). More information about the IDACI index can be found here: https://www.gov.uk/government/publications/english-indices-of-deprivation-2010 .

[57] The principal reason for this methodological choice was that our analysis (at a pre-modelling stage) identified strong curvilinear relationships linking IDACI to IMD as well as to other domain-specific deprivation indices considered (e.g. education deprivation; health deprivation; crime deprivation). This observation confirmed that variability in respondents' IDACI scores should be expected to reflect variability in deprivation levels experienced by respondents across multiple domains. At the same time, IDACI is by definition focused on young people (aged 15 or under) living with income-deprived families. It is therefore particularly relevant, from a conceptual perspective, to the research questions that this study explores.

highest IDACI quintile, 77.8% of young people achieved this threshold, compared to 45.7% of young people in the lowest IDACI quintile (a difference of 32%). As such, when considering each factor in isolation, the relationship between IDACI and attainment appears to be much stronger than that of region and attainment.

Figure 21 Attainment (total capped GCSE or equivalent points) by IDACI quintile
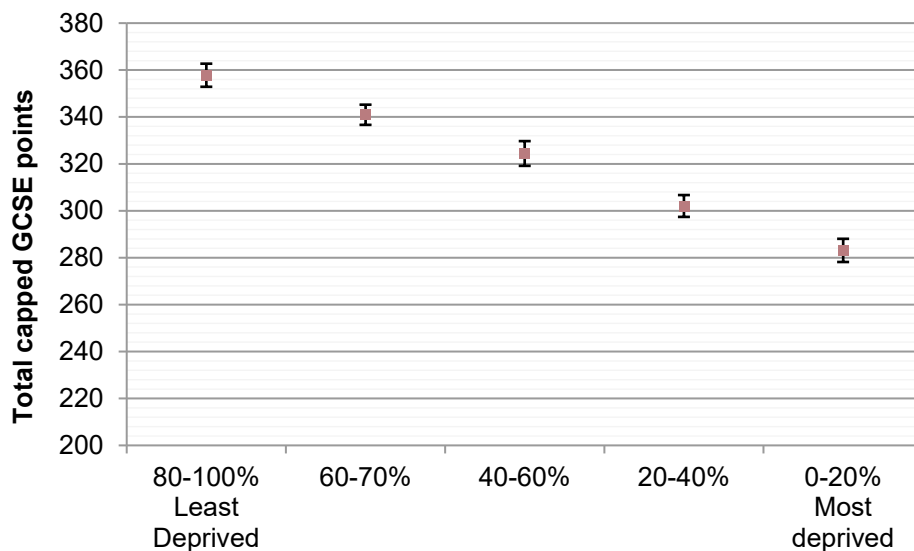
Other intermediary factors which may explain attainment

So far in this chapter, we have focused on the key characteristics of the young people themselves, their schools and the areas they live in and have described the way that attainment clearly varies, most notably highlighting the extent to which young people living in less advantaged circumstances achieve lower grades on average than their less disadvantaged peers. Each of these topics includes factors that are strongly associated with a young person's achievements in school, and together drive a substantial proportion of the variation in attainment. They do not, however, determine young people's outcomes directly or absolutely. Disadvantage is mediated through a variety of channels (such as lack of space or encouragement to study) and consequently some young people with similar characteristics and in similar circumstances will do better or worse than others. Factors such as attendance at pre-school, for example, have been shown to correspond to later attainment differentials (see Taggart et all, 2014) [58]. A host of factors are important in understanding attainment at KS4, many of which may "mediate" the impact of, for example, gender or family income. These factors can optimise the talents of a student, or let them down.

---

[58]

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/455670/RB455_Effective_pre-school_primary_and_secondary_education_project.pdf.pdf

In Chapter 3, we holistically take account of all the structural variables described in detail above, and also consider four further domains which we summarise in the orange boxes in Figure 22 below. These additional domains consist of factors relating to the young person's home environment, their attitudes and behaviours, their health and wellbeing, and the influence of their peers and teachers. Together, these help us to understand attainment in terms of transmitting or ameliorating advantages and disadvantages. Some of these can be influenced or changed through social policy and so are of particular interest. We now describe each of these four areas briefly to give an impression of the range of additional factors we take into account in Chapter 3.

Figure 22 Additional intermediary domains affecting young people at school and individual level (in orange)
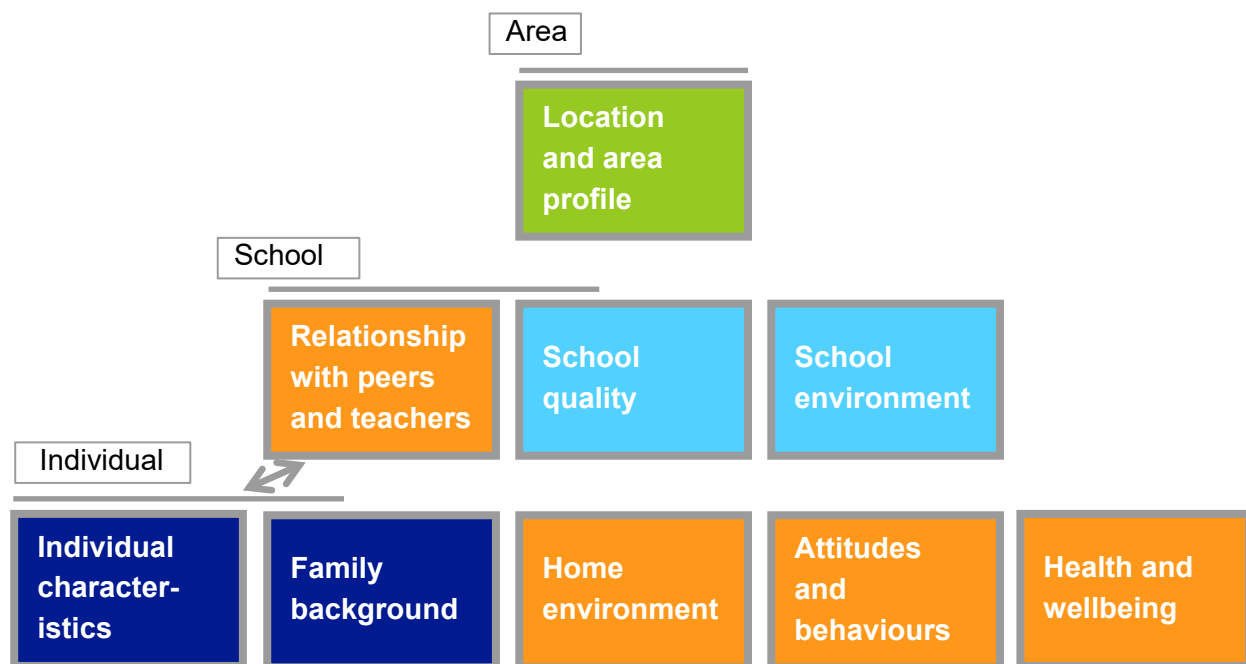
Home environment

The young person's home environment is an important and extensive domain, itself measured here using four sub-domains.

- The first of these is **parental aspiration and expectations**. Here we have focused on what the parent would like the young person to do after year 11 and whether they expect their child will go to university. We found that both of these were linked to attainment, with attainment higher amongst young people with parents who wanted their child to stay in education after Year 11 and among parents who were more likely to think that their child would go to university.
- The second relates to **family resources and the material support** that the home environment provides for the child's education. Here we include whether the young person had access to an internet-connected PC or laptop, tablet or smartphone and whether they received extra tuition in core or non-core subjects. Again, as

might be expected, these were also found to relate to attainment, with average KS4 scores higher amongst those young people with greater resources available to them.

- The third relates to **parental engagement in the child's education**, for example how regularly they read school reports, whether they attended parents' evenings, whether they felt involved in their child's school life, whether they helped with homework, and whether they talked with the young person about their future study or the day's events at school. These factors were all linked with attainment, but for two of these, the direction of the association was not what might be expected. Firstly, young people whose parents felt 'fairly' or 'not very' involved in their school life had higher attainment than those whose parents felt 'very' involved. Secondly, attainment was higher amongst young people in households where parents did not help them with homework compared to those whose parents did. One possible explanation might be that parents are, by necessity, very involved when their child has special needs or is unable to engage effectively with school; young people who are more independent (or able to be independent) are likely to do better in school.

- The final aspect of the home environment is the **quality of the parent-child relationship** which we considered using measures of family cohesion (whether they have family meals together) and family closeness (how well the parent gets on with the young person and how often they argue). Here we find that the 'lower the quality' of the parent-child relationship, the poorer the attainment of the young person.

As summarised above, all of these factors were associated with attainment, but the measures which were particularly pronounced were parental aspirations and expectations as to whether or not their child will go to university, and whether the young person had access to a home computer or laptop with internet access.

Attitudes and behaviours

We also took account of a very large number of measures of young people's attitudes and behaviours. These can be categorised into a number of different sub-domains, as below.

- **Aspirations and expectations**: included measures of what the young person hopes to be doing after year 11; the likelihood of them applying for university; and their expectations of getting in if they did apply. All of these were associated with attainment, with mean scores much lower amongst young people who did not expect or aspire to stay on to study or to go to university.

- **The young person's self-concept**: included two latent concepts[59]; locus of control and the extent to which young people equate hard work with success, both of

---

[59] These were identified and validated using confirmatory factor analysis (fit statistics: RMSEA: 0.030; CFI: 0.994). Further details can be found in a previously published report, Lessof. C, Ross, A., Brind, R., Bell E.

which were measured on an eight-point scale. In addition, we also examined measures of ability belief (the young person's assessment of their own school work and their subjective report of their teachers' assessment of their school work). Taken in turn, all of these variables were strongly associated with attainment. For example, each one unit increase in our 'locus of control' scale, and the scale measuring the extent to which the young person equates hard work with success, was associated with higher attainment, on average 13.9 and 13.5 points. Ability belief was also strongly linked with attainment, with young people who rated their own school work as 'very good' or 'above average' having higher attainment than those who rated it as 'average'. Similarly those who rated their ability as 'below average or not at all good' had poorer attainment. The same pattern was evident in terms of the young person's perception of their teacher's rating of their work;

- **Risky behaviours**: included measures of whether the young person had: drunk alcohol two to three times a month or more in the last 12 months; was a current smoker; had tried cannabis; engaged in petty crime in the last 12 months (either graffitiing, damaging property or shop-lifting); and engaged in fighting (with or without weapons) in the last 12 months. Attainment was lower amongst young people who engaged in each of these risky behaviours, except drinking alcohol where the difference was not statistically significant[60].

- **School behaviours** included measurements of: whether and how much the young person truanted; whether and how much they misbehaved in class; and whether they complied with homework (in terms of doing all, most, some or none of the work set). We also included a summary of attitudes to school. There were some very strong associations here with attainment. In particular, we saw much lower attainment among those who had significant periods of truancy and those who misbehaved in the majority or all of their classes. In addition, we saw a strong positive relationship between attitudes to school (measured on a scale of 1 to 9) and mean attainment - each additional point in positive attitude was associated with a 6 point higher average attainment (equivalent to one grade).

- Finally, we included a number of measures which try to measure how engaged young people are with social, cultural, sporting and community-based activities. These included whether in the last four weeks or 12 months prior to the interview they had: played sport; attended a football match or other sports event; played a musical instrument; participated in community work; attended a youth club (or similar); been to the cinema, theatre or a concert; and played snooker, darts or pool. There were also measures of whether the young person had been to classes

and Newton, S. (2016) Longitudinal Study of Young People in England cohort 2: health and wellbeing at Wave 2, DfE Research Report 501, henceforth referred to as LSYPE2 RR2: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/599871/LSYPE2_w2-research_report.pdf
[60] Appendix B includes tables which show the relationship of each risky behaviour with the two outcome variables. However, readers need to consult Appendix D to understand whether and how much each of these helps explain attainment and progress.

or courses connected to a religious establishment and of their frequency of reading for pleasure. With the exception of attending a youth club and playing snooker, darts or pool, there was a positive association between taking part in all of these activities and attainment. Playing a musical instrument and participating in community work had the strongest face-value relationship with attainment (though it is again important to bear in mind that we are looking at these factors in isolation in this chapter and a slightly different pattern emerges in Chapter 3).

Health and wellbeing

A major focus of LSYPE2 RR2, was the health and wellbeing of young people in year 10 in the first and second cohorts of LSYPE. In the current report we have replicated some of the summary measures that were the focus of LSYPE2 RR2 to examine their link with attainment. These include:

- The **young person's general health** (very good, fairly good, not very good, not good at all). Here we found that those who rated their health over the last 12 months as being 'not very' or 'not at all good' had poorer attainment than those who rated their health as 'very good'.
- The young person's level of **psychological distress** based on responses to the General Health Questionnaire 12 (GHQ12). More information about the derivation of this measure can be found in LSYPE2 RR2 on pages 62-63[61]. The relationship between attainment and psychological distress is highly significant and "curvilinear". Very low levels of psychological distress were associated with lower levels of attainment on average, attainment then increased as psychological distress increased up to a certain level, after which attainment then fell as psychological distress reached higher levels. There is further discussion of this interesting relationship in Box 3 on page 87 **Error! Bookmark not defined.**of the current report.
- **Amount of sleep on a school night**. This was calculated from a set of questions about the young persons' typical sleeping habits during the past month. Using this data, we calculated a continuous variable capturing average amount of sleep on a school night, correcting implausible or missing data items where possible. Inevitably averaged self-reports are an approximation, but they provide a useful indication nevertheless. The recommended guidance suggests that young people in this age group should be sleeping for around 9 hours each night[62]. In relation to attainment we found that those who had 9.5 or more hours of sleep a night had

[61] Lessof. C, Ross, A., Brind, R., Bell E. and Newton, S. (2016) Longitudinal Study of Young People in England cohort 2: health and wellbeing at Wave 2, DfE Research Report 501 can be found at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/599871/LSYPE2_w2-research_report.pdf

[62] http://www.nhs.uk/Livewell/Childrenssleep/Pages/howmuchsleep.aspx

poorer achievement than those who had closer to the optimal amount of sleep or slightly less (8 hours to less than 9.5 hours).

- For the purpose of this report we also examined **bullying**. This analysis incorporated different types of bullying including: name calling; social exclusion; being threatened with physical violence; actually being hit; money or things being taken away; and cyber-bullying[63]. Young people who had experienced any kind of bullying achieved lower levels of attainment on average than those who had not. This was most pronounced amongst those who had been bullied for money or personal possessions, and was least pronounced among those who had been called hurtful names or excluded from a group of friends or from joining in activities.

Subjective experience of school

Although the key structural and quality measures of the pupils' schools were addressed earlier in this chapter (see page 41), we also took account of some additional subjective factors which relate to young people's experiences within school. These covered two main areas.

The first was the attitude of the young person's peers based on their level of agreement/disagreement to four statements: friends think doing well in school is important; friends laugh at those who do well in school; friends distract me from doing well in school; and friends help me with school work. All of these had a strong and significant relationship with attainment - attitudes which support working hard at school were associated with higher attainment scores.

The second area was the relationship of the young person with their teachers: whether they were praised by their teacher; the proportion of their teachers they liked; whether their teachers made the young person work hard; whether teachers cared if they work; whether the teacher picked on them more, the same or less than others; and whether teachers kept order in class. All of these were found to have strong and significant relationships with attainment - more positive relationships were associated with higher scores.

The benefit of having intermediary measures

Almost all of the measures described above in these four intermediary categories have a significant relationship to attainment in themselves; indeed they were selected because we expected them to predict attainment. Although our main interest in this report is in identifying the primary predictors of attainment discussed at the start of this chapter, it is also valuable to consider *how* differences in attainment across the primary characteristics

---

[63] This differs slightly from the approach to measuring bullying in LSYPE2 RR2 as it includes cyber-bullying which was not addressed in the first LSYPE cohort.

can be explained through differences in material resources, attitudes, behaviours, health and wellbeing, school experiences and so on.

Investigating these 'mediating' variables is important, in part, because it can help us to understand how the effects of disadvantage operate and, as a result, identify suitable policies to ameliorate such effects. In this report, we explore some of the ways that intermediary variables can help us understand how differences in attainment among young people from different backgrounds operate. However, LSYPE2 offers far greater opportunity to explore this in more depth than we are able to do here.

It is also important not to place too much emphasis on the relatively simple bi-variate analysis in this chapter. For example, we find that regular parental attendance at parents' evenings is associated with higher attainment. As such, it is possible that persuading more parents to attend parents' evenings regularly will raise attainment as parents gather information that helps them support their child. However it is quite likely that attending a parents' evening is, in fact, an indicator of a parent who is generally supportive of education and who reinforces key messages from school. It is also possible that not attending parents' evening may, for example, be indicative in some cases of a severe breakdown in the parent-child relationship or a very time-pressured working parent who is otherwise struggling. As such, increasing attendance at parents' evenings will not necessarily result in higher attainment if the positive effects of this activity do not follow. While controlling for these other potentially related measures, as we do in Chapters 3 and 4, cannot strictly establish a causal relationship, it does allow us to better understand how different factors interrelate and correspond to attainment.

In summary, even quite basic assumptions need further exploration and we should not assume uni-directional causality. This example is merely illustrative of how we might want to look at these 'intermediary' factors in understanding variations in attainment.

In this chapter we have looked at a very large number of measures in turn. In the next (Chapter 3) we build complex statistical models containing a large number of these variables which makes it possible to see which of these measures help to predict attainment when everything else remains equal. There is also some further investigation of the mediating factors discussed above. Finally, in Chapter 4 we carry out additional statistical modelling, but instead of looking at *attainment* we consider *progress* from key stage 2 (KS2) to key stage 4 (in other words, during the first five years of secondary school).

Chapter 3: How much of the variation in attainment is explained by pupil, school or area factors

Introduction

In Chapter 2 we looked at two measures of pupil attainment at KS4 (total capped GCSE or equivalent points score and the achievement of the Level 2 English and maths threshold) and examined how these outcomes varied across the pupil population. We considered a large number of aspects of young people's lives across seven domains: their personal characteristics; family background; home environment; attitudes and behaviours; health and wellbeing; the characteristics of the schools in which they studied; and characteristics of the areas in which they lived. We looked at each aspect in turn, comparing levels of attainment across different sub-groups of young people. Throughout Chapter 2, we stressed the need to consider these characteristics collectively in order to identify which measures best explained variations in attainment. This is the task we turn to in this chapter.

Technical issues involved in modelling

There are five technical issues that need to be explained before examining the data in this way. We try explain these simply, but those who are less concerned with understanding **how** the analyses were conducted should skip to the section headed "Overall, which factors were the key predictors of attainment?" on page 62. Further detail on our approach is also included in Appendix A Methodology on page 119.

Selecting the outcome variable

The first point to note is that our focus in this chapter is primarily on the outcome variable Best 8 score because it is a continuous measure and therefore provides a stronger basis for complex statistical analysis than Level 2 English and maths threshold which is a binary measure which only indicates whether a pupil is above or below a particular threshold. We briefly summarise the findings from the Level 2 English and maths threshold analysis at the end of this chapter, focusing on similarities and differences with the results for Best 8 points scores.

Accounting for clustering

Secondly, we know that pupils are clustered within schools, and schools are clustered within areas. Indeed, one of the strengths of LSYPE is that young people are sampled to reflect this structure, as shown in Figure 23 below.

Figure 23 Young people are clustered in schools which are clustered in areas

Area

School

Pupil
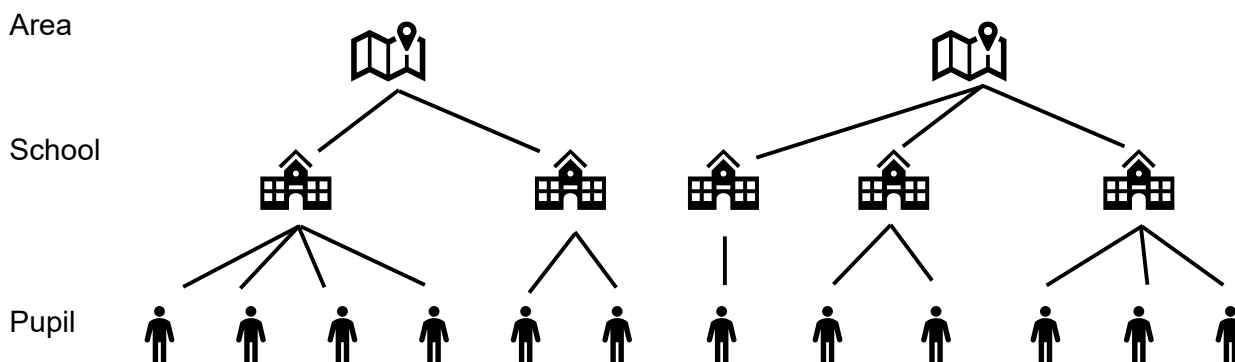
The structure of the dataset therefore has particular benefits - as we explain later, it can help us identify 'school' and 'area' effects in addition to individual effects. However, a logical consequence is that our analysis needs to take account of the fact that pupils selected within any given school are more likely to be similar to each other than to pupils from other schools, because of their shared experience of the same school environment. The same principle may apply to schools being more alike to other schools within a particular area than to schools outside that area. This lack of independence between sample members breaches a fundamental assumption of many basic statistical modelling approaches. We therefore use a more complex approach – multilevel modelling – to account for this clustering of pupils (the first level) within schools (the second level) and of schools within area (the third level). Furthermore, this approach allocates the proportion of variation in pupil attainment that is attributable to pupils, the schools they attend, and the areas in which their schools reside. This helps us identify how far differences in pupil attainment are to do with pupils themselves, the schools that they attend, or the areas in which their schools are situated.

When we subsequently add information into our models about the characteristics of each of the three levels to identify the important predictors of attainment, we will also identify how much each of these measures contributes to explaining variation at the three levels of pupil, school and area. For example, it is well established that differences in attainment between schools has a lot to do with the kinds of pupils attending them[64]. Therefore, when we include information relating to the pupils, such as their parents' social class, education and income, we find that we can explain a significant amount of the variation in attainment *between schools* as well as between pupils.

---

[64] For example, research from the Sutton Trust highlighted difference in attainment according to the deprivation levels of the school intake. See: Attainment gaps between pupils in the most deprived and advantaged schools, Noden, 2009
https://www.suttontrust.com/wp-content/uploads/2009/05/Attainment_deprived_schools_full-1.pdf

By extension, through modelling the data in this way, we gain a better understanding of the extent to which differences in attainment between schools are to do with differences in pupil intake, and the extent to which they relate to 'true' differences between schools. Finally, as well as providing a better understanding of the key factors that explain attainment and the level in which they operate (pupil, school, area), we also get a sense of how much is left *unexplained* at each level. This serves as an indicator of how successful we have been at understanding the phenomenon of interest – in this case 'attainment'.

Exclusion of individuals from the model

Unfortunately, as with most social surveys, many individuals are missing data for one or more variables. A widely accepted approach is to carry out statistical modelling using a sub-sample of 'completed cases' where all the necessary data is available. Since this study uses a substantial number of measures, the level of missing-ness is large – from the total sample of 8,890 pupils in LSYPE who are eligible for this analysis[65], the number with complete data is 4,200. In Appendix A, we examine and discuss the extent to which this might bias the findings presented here. Briefly, a comparison of those who remained in the sample with those who were excluded shows that our completed cases sample is more advantaged than the general population. However, close comparison of the effects of our measures across more inclusive samples suggest the impact on our findings is relatively small. Nevertheless, without multiply imputing data, we cannot be absolutely certain of the true impact. It is most likely that we will have underestimated some of the effect of disadvantage and some of the factors associated with this. In terms of attainment, the mean attainment score for the completed cases sub-sample is 348.2 (SD=70.6) compared to the 319.4 (SD=94.5) for the full sample reported in Chapter 2 – this equates to a difference of around 5 GCSE grades[66]. In other words, we have retained young people with higher and less dispersed scores. Similarly, the completed cases sample has a higher rate of Level 2 English and maths attainment with 72.8% achieving the threshold compared to 61.7% in the total sample reported in Chapter 2[67,68]. A further related point is that our analysis presented here relates to mainstream schools as we had insufficient information about individuals in special schools of different kinds.

---

[65] There were 9,076 young people who responded to Wave 3 and consented to data linkage, 9,035 for whom a unique reference number identifying their school was available in the dataset and 8,890 who were not attending a special school of some kind. Unique school identifiers were missing for a total of 41 Wave 3 respondents who had consented to data linkage and 145 students attended special schools.

[66] The mean for the 8,890 with a known school that was not a special school was 324.9 (SD=86.5).

[67] This figure includes young people for whom we have no school information and those who were in special schools. If we exclude these individuals, 63% of the remaining 8,890 individuals achieved Level 2 basics.

[68] We noted earlier that the LSYPE2 sample already has higher attainment than the national average (61.7% compared to 59.2%) suggesting there are certain kinds of individuals with lower attainment outcomes already missing prior to individual item missingness.

Findings for small sub-categories

A closely related issue is that whilst the overall LSYPE2 sample is very large (even taking account of the loss of cases with missing data), some of the sub-categories for measures in our analysis are quite small. Our confidence in the estimates for these categories will be lower than for categories where we have larger sample sizes. In Appendix B we show the sample sizes across all of the measures in our analysis and in Appendix D the full analysis models include confidence intervals for the estimates quoted in the report. In basic terms, these confidence intervals indicate the range within which you can be 95% certain that the true value lies. Findings which are based on a sub-category with a very small sample size will tend to have large confidence intervals, meaning that even quite large differences in attainment may not be statistically significant. Throughout the report, we focus our commentary on those differences which are statistically significant. In cases where the sub-category sample size is low, genuine differences may not meet the threshold for statistical significance and are not therefore generally discussed. In cases where differences relating to small sub-categories are flagged in the text, the difference is statistically significant unless otherwise specified, but it should be noted that the exact scale of the difference will be subject to less certainty (i.e. the data will have wider confidence intervals) than would be the case for a sub-category with a larger sample size.

Causation

Our final caveat is that, throughout this chapter, we describe the way that some differences in attainment (such as girls doing better than boys) may be "explained" by differences in behavioural or attitudinal measures which differ between the genders (such as parental aspirations for post year 11). We use the term "explain" in a statistical sense. For example, in the case above, we would mean the extent to which a given behavioural or attitudinal measure may account for some, or all, of the association identified between gender and attainment. Whilst it is true that this factor might also explain why girls have higher levels of attainment than boys in a real sense of the word, we often cannot infer this level of causality. To do that, we would need clear evidence about how the causal relationship actually works, which is not something we can easily do with an observational study of this kind or with this statistical approach. A finding might support a particular hypothesis but it is not its absolute proof.

How much variation is at an individual, school or area level?

We begin with the simplest possible multilevel model (often referred to as the base model) which includes our measure of attainment and three indicators, identifying each of the three levels of pupil, school and area. The purpose of this is to "decompose" the

variance[69] in attainment between these three levels - in other words to see their relative importance in explaining differences in attainment.

We found that each was statistically significant in explaining attainment; approximately 84% of the variation in attainment was attributable to the pupil level, 16% to the school level[70] and a negligible amount (less than 1%) was attributable to the area level. Observing a trivially small (but significant) area effect is not uncommon in studies of this kind[71] and is likely to do with the fact that the indicator of locality that we used (Local Authority or LA) is quite heterogeneous, and LAs are insufficiently different from one another. Given that the computation for estimating three level models was excessively large we decided to proceed with just two levels; individual and school, although we continued to take account of the area in which the young person lived by including individual level variables within the model, specifically the region, level of IDACI for their home address, and whether this was an urban or rural setting[72].

Having established that 84% of the variation in attainment was attributable to pupils and 16% to schools, we then attempted to explain the variation in attainment by adding sets of covariates (the characteristics in each of the domains described in Chapter 2) into the model. Table 1 below shows how much variation in attainment was explained as we included each additional domain. It is important to note that this information is indicative and will depend on the order in which the domains were introduced i.e. each additional domain was only able to explain variation that had not already been explained by a prior domain. It is also worth reiterating a point made above. Variance in attainment attributable to the school level is also a consequence of the school's intake. This is why a significant amount of school level variance is accounted for by measures associated with the individual (e.g. personal characteristics, family background, and home environment). Furthermore, measures for school also include measures of school composition such as percentage of ethnic minority pupils, percentage eligible for free school meals.

Overall, we can explain around 57% of the variation attributed to the pupil and 87% of the variation attributed to schools.

---

[69] Variance and variance decomposition are the proper statistical terms that should be used here, but we use the more familiar word 'variation' throughout the remaining sections of this report
[70] These figures are based on the analytic sample of 4,200 cases for whom we have all necessary measures. The variance partitioning of the full sample excluding individuals for whom school was missing or who attended special schools (8,890 individuals) was almost identical, with 83% variation at individual level and 17% at school level. It is important to acknowledge that had we not excluded these cases, a far greater percentage of the variation in attainment in schools (approximately 38%) would have been attributable to the school level. This reflects the fact that Special schools had much lower levels of attainment on average, driving up the variation in scores overall.
[71] This was also found to be the case for GCSE attainment in 2003 with less than 1% variance observed at LA level, see DFES. (2004). *Statistics of Education 02/04: Variations in pupil progress 2003*. London: Department for Education and Skills
[72] When we had agreed our final models using two-levels, we re-ran the analysis using a three-level approach to see whether the findings were sensitive to this decision. The results were remarkably similar except in a few instances where a measure was previously borderline significant and tipped in to non-significance or vice-versa.

Details of the variables included in each domain can be found in Appendix D.

Table 1 Additional variation explained by each domain as it is introduced into the model

Domain introduced to the base model	Individual level	School level
Personal characteristics	18.1%	10.5%
Family background	12.1%	38.1%
Home environment	14.5%	15.6%
Attitudes and behaviours	10.5%	6.4%
Wellbeing	1.2%	0.1%
School[73]	0.6%	14.6%
Area[74]	0.2%	1.3%
% of total variation explained	57.1%	86.6%

Overall, which factors were the key predictors of attainment?

This statistical modelling showed that a remarkably large number of measures were important in predicting attainment. In this section, we briefly summarize which of our measures were significant in the **final** model, when all measures were taken into account simultaneously.

Understanding the data

Figure 24 on page 66 and on Figure 25 on page 68 provide a highly simplified visual summary of the final model to give a sense of the relative importance of different measures. These figures only include measures which had a significant association with attainment and only include those sub-categories that were significantly different from the reference or comparison category (for example the effect of truanting for several days at a time compared to not truanting at all).

Figure 24 focuses on the stronger associations and only includes those categories for which there is a difference equivalent to at least two grades (12 points) between the

---

[73] In addition to measures that were measured at the school level (e.g. school composition and school quality), the school domain also includes individual level measures that capture young people's subjective perception of their peers and their teachers

[74] All measures relating to area were measured at the individual level, based on the postcode where the young person lives.

reference category and at least one of the related categories. The categories have been re-ordered so that the categories with the lowest attainment appear first, though it should be remembered that there are margins of error around all of these statistics, so this ordering is indicative only. Most of the bars lie to the left of the vertical axis and are the characteristics associated with lower attainment, relative to the reference category. Those to the right of the vertical axis, particularly in Figure 25 are associated with higher attainment. In some respects, the distinction between positive and negative is arbitrary (we could transform the way we describe some measures so they all fall on the same side, but this would add a layer of unnecessary complexity). The important issue is really the *strength* of the relationship, rather than the sign. In Appendix D we provide the models in full and the 95% confidence intervals around each of these estimates. These are not shown in the graphs in the interests of legibility.

In the remaining part of this section we give an overview of the factors which have the strongest relationship with attainment after controlling for all the other factors in the model. Then, in later sections of the chapter we move on to look at the more detailed models and consider groups of measures within their respective domains, which helps us to explore some of the initial findings in more depth.

Predictors of low attainment

To begin with, several factors shown towards the top of Figure 24 were associated with very much lower levels of attainment, equating to about five GCSE or equivalent grades lower per student (differences of 30 points or more). Some of these categories had extremely small sample sizes and so need to be treated with caution (again, see Appendix D for details of the confidence intervals associated with these data).

The most striking measure is **truanting** – the tiny proportion of young people who truanted a lot achieved much poorer results even after controlling for all the other factors in our model. Pupils who reported truanting for weeks at a time had far lower levels of attainment than those who did not truant (-104.5 points) as did those who truanted even a few days at a time (-61.5) or who refused to report how much they truanted (-39.0). The most obvious interpretation is that these students did badly in exams because they had missed a lot of school, but the direction of causality may be reversed (i.e. missing school may have been a consequence of doing badly in school and disengaging) and a number of other interpretations of this relationship are possible.

Levels of attainment were also much lower where a young person thought they would be **likely to take up full-time work or do something else** after year 11 (-50.4 points lower) though this affected only a handful of young people (0.2% of the sample). Similarly, attainment was much lower where a parent believed it was **not at all likely their child would go to university** (-48.3 points), or that they were not very likely to (-29.3), were fairly likely to (-15.0) or the parent was uncertain (-20.9). There may be a number of effects going on here. Students who are less academic or who are already struggling to

engage with school may be reported as less likely to go to university by their parents. At the same time, students who are not expected to go to university by their parents may not be motivated to work hard towards achieving higher attainment if university is not seen as either attainable or relevant.

Much lower levels of attainment were associated with being identified as **having SEN** (including statements and/or school action) (-34.7 points lower) and with having no biological parents living in the household (-33.5 points lower). Again, it was to be expected that attainment for those with SEN would be lower, given the challenges which they must overcome. Another personal characteristic strongly associated with attainment is **having a disability that affects school** (-14.2). Even though our model takes account of a very broad range of factors that might mediate any independent effect of ethnicity, both **Black Caribbean and Black African students** had significantly lower attainment than White students in similar circumstances (-12.4 and -11.0 points respectively, equivalent to about two grades).

The young person's own **attitudes and behaviours** were associated with up to three grades lower attainment. For example, young people who were undecided about university or did not know if they would apply had lower scores (-21.9) than those who reported they were very likely to go to university. Perhaps surprisingly, this difference in attainment was more pronounced than for those who reported that they were not at all likely to go to university (-18.6) or were only fairly likely to go (-5.8). Confidence in ability to get into university also mattered. For example, believing they were not very or not at all likely to be successful if they did apply was associated with scores 20.4 points lower than those of young people who believed they were very likely to get in. The young person's assessment of their own schoolwork was less important, but nevertheless played a role in explaining attainment, over and above these measures, with young people doing better if they thought their school work was very good (+12.4) or above average (+7.8) and worse if they thought it was below average (-12.3). Similarly, those who believed that their teachers would assess their work as above average attained an average of 12.1 points more than those who reported that teachers assessed their work as being average. Those who thought their teacher would assess them as 'very good' attained an average of 14.4 points more than those with average assessments. To some extent this may be a self-fulfilling finding – it is not necessarily surprising that those who receive less positive feedback on their work display lower attainment. However, there may also be more complex factors at play, for example the possibility that negative feedback is demotivating and negatively affects attainment.

In addition to truanting, another important behaviour related to **homework**. Young people who did all their homework did better than those who reported doing most of it (-10.3 points) or some or none of it (-21.7). Interestingly, another behaviour that was strongly predictive was the frequency with which the young person reports **reading for pleasure**.

Those who did not read for pleasure on most days attained significantly lower grades than their counterparts, most noticeably if they never read (-14.0 points).

One school measure was found to be amongst the most important predictors of lower attainment. Attending a **non-selective school** was associated with just over a three-grade disadvantage (-20.1) compared to attendance at a selective school. Again it is perhaps worth emphasising that all findings throughout this chapter are based on outputs from our model, which controls for other factors such as the family background of young people attending a given school. The absolute difference in attainment between young people at selective and non-selective schools, as discussed in Chapter 2, was much larger than this. It is also worth noting that the analyses in this chapter do not take account of prior attainment which is highly important when considering selective schools – we discuss this further in Chapter 4.

Four measures which captured the **social and economic position** of the young person were among the most important determinants of attainment. These are described in more detail in the next section but, in brief, those who lived in owned or mortgaged housing had higher attainment than those living in Housing Association rented (-10.5 points lower) or private rented housing (-16.1). Those who lived in a household with above median income had higher attainment that those living in a below median income household, particularly one where the young person was eligible for FSM (-12.2 points) or had been eligible for FSM within the last 6 years (-11.4).

Region, was also important, with the North West, for example, being associated with 13.1 points lower attainment than was achieved by young people in London.

Figure 24 Summary of the final model (significant effect of two grades or more between reference category and one or more comparator categories)

Category	Value
[REF Not at all] …truants weeks at a time	-104.5
…truants several days at a time	-61.5
…truants but DK or refuses amount	-39.0
[REF Stay in educ] …work full-time/something else	-50.4
[REF P thinks YP v.likely to go to Uni] ...not at all likely	-48.3
…not very likely	-29.3
…doesn't know if YP will go to Uni	-20.9
…fairly likely to go to Uni	-15.0
[REF without SEN] … With Special Education Needs	-34.7
[REF Two bio. parents in HH] …no bio. parent in HH	-33.5
[REF YP very likely will apply to Uni] ...DK/undecided	-21.9
…not at all likely to apply	-18.6
…fairly likely to go to apply	-5.8
[REF Usually does all homework] ...some or none	-21.7
…does most	-10.3
[REF YP v.likely to get in if applies] …not v./not at all…	-20.4
...DK/undecided/will leave full-time education	-15.1
…fairly likely to get into Uni if applies	-7.3
[REF Attends selective school] ...non-selective school	-20.1
[REF Mother has a degree] ...Other qualifications	-17.7
...No quals	-15.3
...Level 1 or equiv.	-11.1
...A Level or equiv.	-8.1
...5+ A-C* GCSE or equiv	-6.8
...Higher Educ but less than degree	-5.5
[REF Own/mortgage] ...Private rent	-16.1
… Rent from Council /Housing Assoc	-10.5
[REF YP reads most days] ...never reads	-14.0
…once a week	-12.0
…hardly ever	-11.2
…less than once a week	-5.4
…more than once a week	-4.6
[REF: No Illness/disability] …ill/disability affects school	-14.2
[REF Lives in London] ...North West	-13.1
...West Midlands	-9.5
...Yorkshire and Humberside	-9.4
…East Midlands	-8.8
…South East	-8.1
[REF White] ...Black Caribbean	-12.4
...Black African	-11.0
[REF Self-ass average sch wk] …below/not at all good	-12.3
…above average	7.8
…very good	12.4
[REF Income> median not FSM] …eligible for FSM	-12.2
…eligible for FSM within last 6 years	-11.4
…not eligible for FSM but<median income	-8.0
[REF Teach-ass average at sch wk] …above average	12.1
…very good	14.4

For the sake of completeness, Figure 25 below, shows all of the remaining categories which displayed significant variation from their comparator categories in terms of attainment. All of these measures saw a difference of less than 12 points (i.e. the equivalent of less than 2 grades). We discuss these in detail in the domain-specific sections which follow.

Figure 25 Summary of the final model (significant effect of less than 2 grades between reference category and comparator categories)

[REF Sleeps 8-9.5 hrs] …9.5 hrs or more	-11.4
[REF Has desk/laptop with internet access] …does not	-11.4
[REF IDACI Best quintile] …worst IDACI area	-10.0
…4th quintile	-7.9
…3rd quintile	-6.1
[REF Has not fought without weapons] …has done so	-9.9
[REF Parent always discuss report] …most of time	-8.4
[REF Not cyber-bullied in last 12 mnth] …has been	-7.7
[REF Parent argues with YP once a wk/hardly/never] …most days	-7.0
…more than once a week	-6.4
[REF Not paid for core tuition] …Paid for core tuition	-5.8
Locus of control	2.4
GHQ12 score (1-36)	2.5
GHQ12 quadratic term	-0.1
(Likert scale) Friends don't laugh at those who do well	2.9
(Likert scale) Friends don't distract YP from doing well	2.0
[REF Teachers make YP workhard] …teachers do not	5.1
[REF Parent v.involved in school life] …not v.involved	5.4
[REF School has a Sixth form] …does not	6.0
[REF YP talks w. fam about future study a lot] …a little	6.1
…quite a lot	4.1
[REF Teachers pick on YP same as everyone] …less	8.0
School value added	8.3
[REF Family meal everyday] …once/twice a week	4.4
…or not at all	8.8
[REF Does not do community work] …does	6.5
[REF Does not play sport] …does	8.6
[REF Does not play musical instrument] …does	9.5

Importance of individual level characteristics

In this section we describe in more detail the individual level characteristics that best predict differences in attainment, considering each domain in turn: personal characteristics; social and economic position; home environment; attitudes and behaviours; health and wellbeing; and subjective school experience. Later, we consider school and area level characteristics.

Personal characteristics

In Chapter 2 we showed that when looking at each of the personal characteristics measures in isolation, attainment varied significantly by gender, term of birth, ethnic group, illness or disability and SEN status (including statements and/or school action). However, when we take account of all of our measures (across all domains) simultaneously, gender (see Box 1) and term of birth were no longer found to be significant predictors of attainment. This does not mean that gender and term of birth have somehow stopped being important; rather, other aspects of young people's lives explain or 'mediate' the underlying effects.

Figure 26 on page 71 shows that the strongest explanatory factor of attainment was SEN; those identified in NPD as having SEN status achieved about 6 grades less than their non-SEN counterparts (-34.7 points) all other factors being equal. The association between SEN and attainment (and, as described in Chapter 4, progress) should not necessarily be interpreted as meaning that teachers are not making appropriate adjustments to their teaching – it may, for example, also be a sign that schools are effectively identifying pupils who struggle with formal schooling. Similarly, pupils who reported an illness or disability which affects their schooling achieved just over two grades less than those reporting no illness or disability (-14.2 points). Two ethnic minority groups had lower levels of attainment on average relative to their White counterparts; Black Caribbean students (-12.4 points) and Black African students (-11.0). The relationship between ethnicity and attainment is complex, however, and described in more detail in Box 2.

Figure 26 Personal characteristics associated with attainment in final model
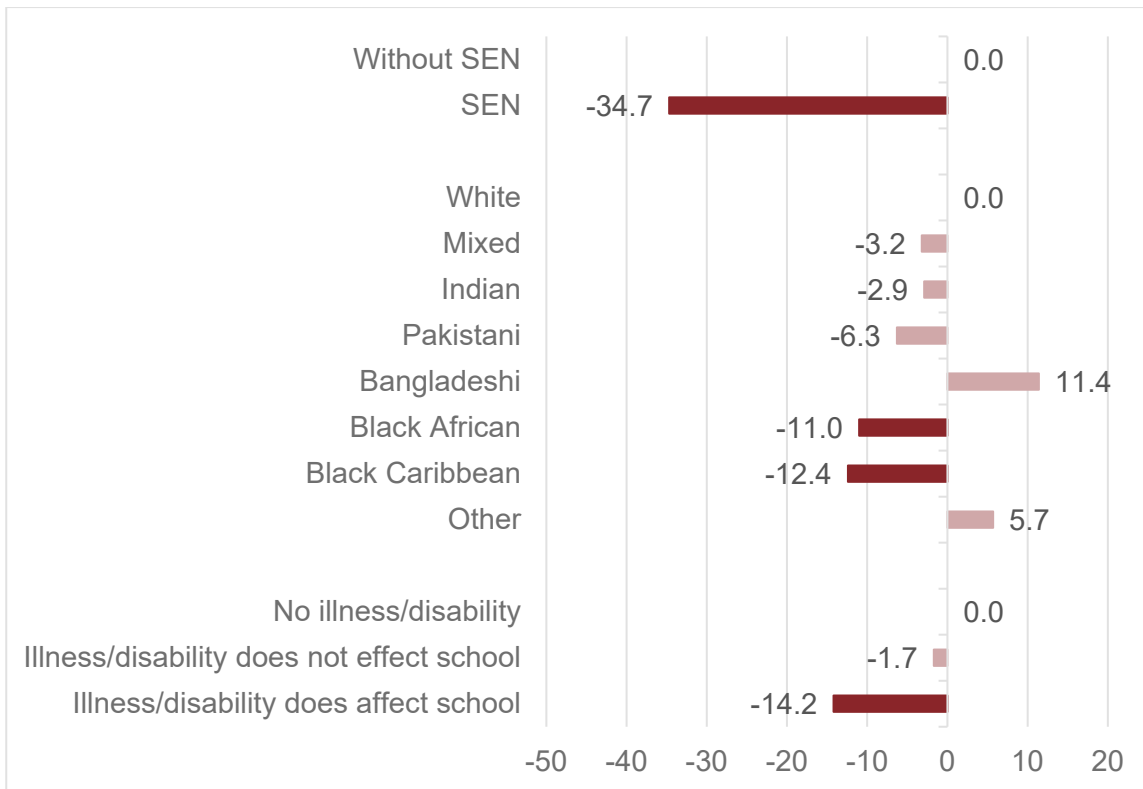
Note: dark bars are statistically significant at a 95% degree of confidence

Box 1 How much of the gender gap can be explained by other measures?

In Chapter 2 we showed that when looking at gender as a predictor of attainment without taking any other factors into account, on average girls' Best 8 scores were 26.6 points higher than those of boys. However, in the final model which took account of all the relevant measures covered by LSYPE, girls only attained 3.2 points more than boys, a difference which is not statistically significant. This does not mean that gender has somehow stopped being important; rather, other aspects of young people's lives explain or 'mediate' the large underlying gender effect. We can investigate this more deeply by looking at how the difference in attainment between girls and boys varies when we start with a very simple statistical model which includes only a single domain and then examine how this changes as we introduce additional domains. The figure below shows pictorially how the difference in attainment between girls and boys reduces as each new domain is added to the model.

Difference in attainment between girls and boys as additional domains are added to the model
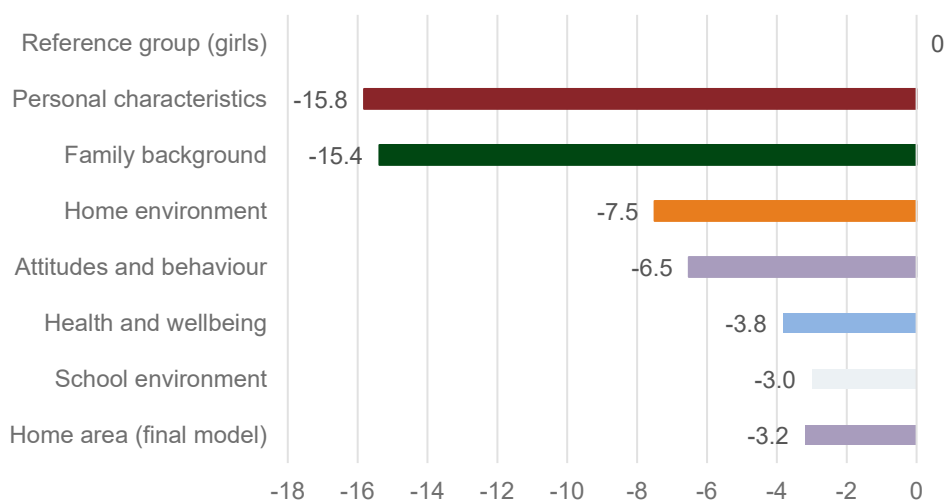
For example, the gender difference first reduces when we take account of other personal characteristics (the red bar above); in fact, the difference falls from 26.6 points when looking at the gender difference in isolation to 15.8 points when personal characteristics are taken into account (a decline of 41%). This is mainly driven by the fact that boys were much more likely to be classed as having SEN (including statements and/or school action) than girls (15.2% of boys compared with 8.7% of girls) [75], and SEN is a strong predicator of attainment[76]. Disability which affects school is also a predictor of attainment and, again, there are more boys than girls in this category.

Following the same approach, we see that the gender gap further reduces as we take account of other domains of a young person's life. The observed difference between

boys and girls almost halves again (reducing to 7.5 points) when we introduce the set of measures which describe young people's home environment (the orange bar in the figure above) alongside personal characteristics. This suggests that boys and girls had different experiences in the home which could account for some of the difference in attainment. Two particular measures stand out in this respect. Parents were much more likely to think their daughters would go to university than their sons. For girls, half of parents thought it was very likely (50.4%) they would go to university while 3.2% thought it was not at all likely. In contrast, for boys, 39.0% of parents thought it was very likely they would go to university and 9.5% thought it was not at all likely. Parental perceptions of the likelihood of going to university were a strong predictor of attainment. Similarly, what the parent wanted the young person to do after year 11 was both strongly predictive of attainment and differed between girls and boys - there was close to a 10 percentage point gap between wanting your child to continue in education (93.4% for girls compared to 83.8% for boys) and a much higher interest in apprenticeships for boys than girls (11.5% compared to 3.8%). Parents also talked with girls more about their future studies, which again had a positive effect on attainment.

The difference between girls and boys reduces further still (to 3.8) when we adjust for all aspects of health and wellbeing (the blue bar in the figure above), in particular reflecting differences in psychological distress where scores differ for girls and boys. When school measures, particularly those related to peer attitudes, teacher relationships and school added value were added to the mix, the difference in attainment associated with 'gender' decreased to the point where it was no longer significant.

By modelling the data in this way, we can begin to unpick the factors that might contribute to differences in attainment associated with a particular attribute or circumstance. To reiterate, this does not mean that gender ceases to be important; it simply means we have begun to understand some of the potential drivers of the difference. Readers who are interested can look at Appendix D and investigate other associations in this way.

Family background

In Chapter 2 we showed that attainment varied by six closely related measures which captured aspects of the young person's social and economic position (their family

---

[75] The higher incidence of SEN among boys is also found in published data on SEN. For example, the National Statistic 'Special educational needs in England: January 2015' shows that 12.7% of boys aged 15-16 and 8.8% of girls aged 15-16 received SEN support.
See: https://www.gov.uk/government/statistics/special-educational-needs-in-england-january-2015
[76] This is one of the instances where losing more girls from our 'completed cases' may have had some effect on our estimates. This analysis, when based on the full sample of girls and boys (including those with missing data) shows that the difference has dropped even further, to 12.4 points. This remains strongly significant.

background). All but two of these measures (number of working parents in the household[77] and NS-SEC), also helped to predict attainment in the final model, as shown in Figure 27 below. In summary, our model shows that even having controlled for other factors:

- Not living with either biological parent had a large negative effect on attainment (-33.5 points compared to those living with both biological parents), most likely reflecting these young people's disrupted lives[78].
- Maternal education was also very important, as other studies have shown[79], perhaps because having an educated mother is likely to contribute to expectations and attitudes conducive to higher attainment and mothers with qualifications may act as aspirational role-models. An educated mother may also be more able to support the young person's education and give practical input, though this should be viewed in light of our later finding that a high level of parental involvement at this stage of education does not necessarily correspond with higher attainment. Pupils whose mother had a degree level qualification achieved higher levels of attainment themselves than other young people. The attainment disadvantage compared to those with a degree educated mother was equivalent to about one GCSE grade in cases where: the mother did not have a degree but had higher education qualifications (-5.5 points); had A levels or equivalent (-8.1); or had a minimum of 5* A-C GCSEs or equivalent (-6.8). The difference increased to approximately a two to three grade disadvantage for those young people whose mothers only held: a Level 1 qualification (-11.1 points); other qualifications (-17.7); or no qualifications (-15.3).

---

[77] As explained in Chapter 2 all parents working means just one working parent if the household has only one parent; this means we treat one parent working in a single parent household as equivalent to two parents working in a two parent household.
[78] The sample of young people living in these circumstances is very small (n=44), and therefore the confidence intervals around this estimate are very wide (CI=-53.0,-13.9).
[79] For example, analyses of the Millennium Cohort Study have identified that higher levels of parental education correspond to higher levels of vocabulary at age 14. See: Initial findings from the Millennium Cohort Study Age 14 Survey - What influences vocabulary?, CLS and IOE, 2017
http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=2419&sitesectiontitle=MCS+Age+14+initial+findings

Figure 27 Measures of social and economic position associated with attainment in final model

Category	Value
Two biological parents	0.0
One biological parent and one step-parent	-1.4
One biological parent	0.1
No biological parents	-33.5
Mother has degree	0.0
Higher Ed.	-5.5
A Level or equiv.	-8.1
5+ A-C* GCSEs or equiv.	-6.8
Level 1 or equiv.	-11.1
Other	-17.7
No quals	-15.3
Mother not present in HH	-8.0
Own/Mortgage	0.0
Rent from Council/Housing Assoc	-10.5
Private rent	-16.1
Other	8.4
HH Income > median and not FSM	0.0
Not eligible for FSM but HH income < median	-8.0
Eligible for FSM within last 6 years	-11.4
Currently eligible for FSM	-12.2

Note: dark bars are statistically significant at a 95% degree of confidence

- Young people living in rented accommodation achieved lower levels of attainment on average than those living in a home that was owned or mortgaged, regardless of whether they rented from the Council or Housing Association (-10.5 points) or the private sector (-16.1). Attainment was also lower if the young person lived in a comparatively poor household. Compared to young people living in households with at least the median income (£25,600 each year), the attainment of young people in households with income below the median but above the FSM threshold was approximately one grade lower on average (-8 points). There was a two grade difference for those who had been eligible for Free School Meals in the past (-11.4) or who were currently eligible for FSM (-12.2). Previous studies point to effects related to a lack of resources, some of which we were unable to measure in the

75

survey. These include factors such as space in which to work, family security and the absence of stress caused by difficult economic circumstances and the inclusion of these in the model may have served to provide further understanding of this effect.

Box 2 The relationship between ethnicity and attainment varies between groups

The relationship between ethnicity and attainment is complex and highlights the varied experiences of different Black, Asian and other minority ethnic groups. It also serves to illustrate how much we can learn by examining changes in predicted attainment when we add additional information on young people's lives in successive stages. Here, we focus on pupil ethnicity but full findings relating to all measures can be found in Appendix D.

When we only include the young people's personal characteristics in our model (dark red bar), we see that several ethnic groups attain higher Best 8 scores than White students, particularly Other (19.7), Indian (20.0) and Bangladeshi (14.2). However, because of a strong association between economic position and ethnic background[80], predicted attainment changes quite significantly when we take account of family background (green bars). The gap in the attainment between Indian and White students reduces slightly (by 3.7 points) when family background is added to the model, explained in part by the relatively advantageous social and economic position of young people in Indian families. In contrast, the average attainment of pupils from BAME groups which were more disadvantaged on average than White pupils tended to increase when family background was added to the model. For example, as a consequence of this adjustment, the difference in attainment compared to White pupils increased to +26.6 points for Bangladeshi pupils (a change of +12.4). Similarly, those from the 'Other' group saw their attainment relative to White pupils increase by a further 6.4 points (to 26.1). Where more disadvantaged groups had lower attainment relative to White pupils, the gap narrowed when family background was taken into account (most notably from -12.6 to -3.0 points for Black Caribbean pupils, which now constituted a non-significant difference).

Difference in attainment between minority ethnic groups and white ethnic group as additional domains are added to the model

---

[80] For example, Bangladeshi families are the most likely to experience economic disadvantage as shown in analysis of Census data from the University of Manchester. See: 'Ethnicity and deprivation in England: How likely are ethnic minorities to live in deprived neighbourhoods?', Jivraj and Khan, 2013 http://hummedia.manchester.ac.uk/institutes/code/briefingsupdated/ethnicity-and-deprivation-in-england-how-likely-are-ethnic-minorities-to-live-in-deprived-neighbourhoods%20(1).pdf

Indian	20.0 / 16.3 / -2.9 / 1.4	
Pakistani	-18.0 / -9.5 / -3.9 / -6.3	
Bangladeshi	14.2 / 26.6 / 6.0 / 11.4	
Black African	-5.0 / 6.5 / -13.0 / -11.0	
Black Caribbean	-12.6 / -3.0 / -14.4 / -12.4	
Other	19.7 / 26.1 / 9.8 / 5.7	

-22   -12   -2   8   18   28   38

■ Personal characteristics ■ Family background ■ Home environment ■ Final model
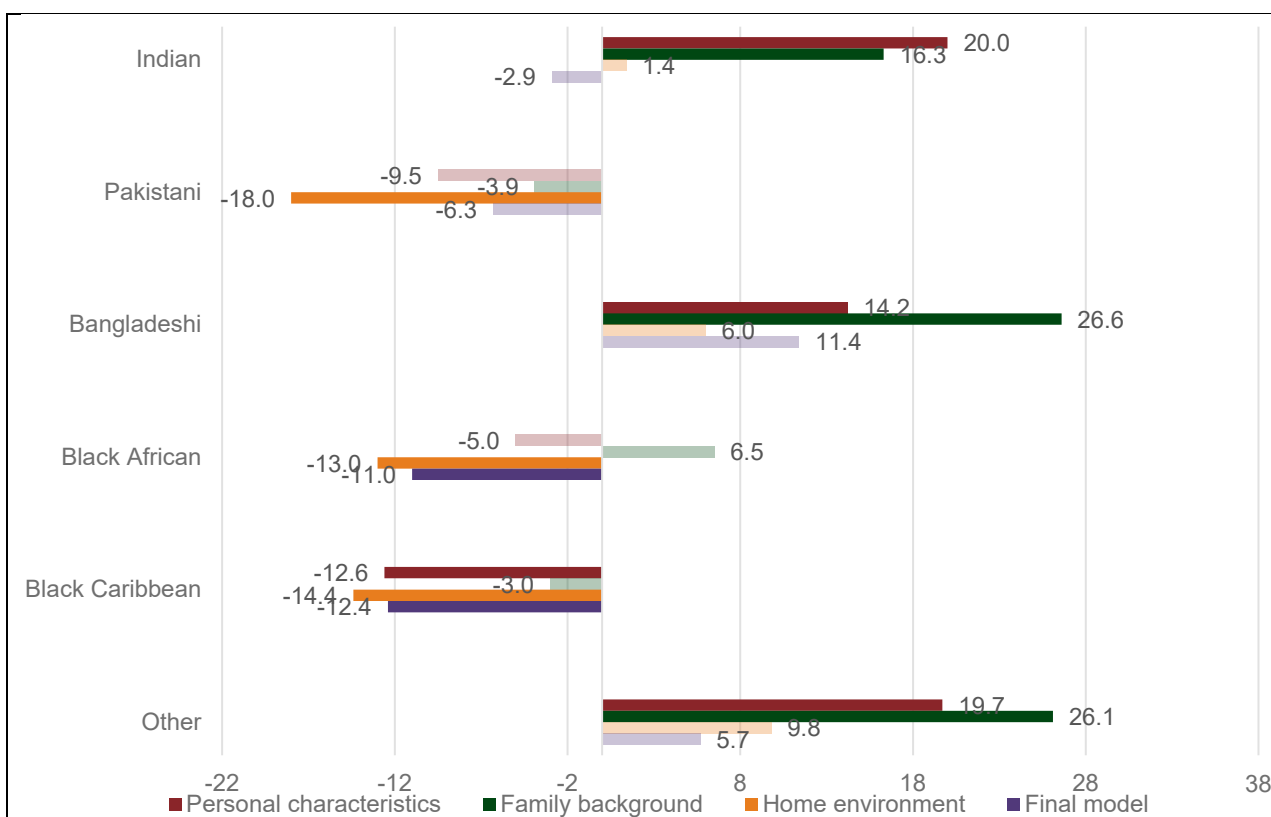
The third, orange bars illustrate the differing impact of another important domain; home environment. All BAME groups had home environments which were, on average, more conducive to educational attainment, for example, having parents with higher aspirations for Further and/or Higher education. As a result of adjusting for these differences, the significantly higher levels of attainment associated with being Indian, Bangladeshi or from the 'Other' BAME category decreases. Similarly, the attainment of Pakistani, Black African and Black Caribbean backgrounds became (more) negative in comparison with White pupils, suggesting that a more attainment-supportive home environment for these BAME groups 'explains' some of their apparent ethnic advantage (or reduces their disadvantage).

Because of the complexity of the data we have not shown the effect of introducing every domain. Nevertheless, similar, although far smaller, changes in the relative differences in attainment can be found when including young people's attitudes and behaviours. Once we take account of differences in aspirations/expectations, attitudes and behaviours across ethnic minorities the relative position of some ethnic groups improves slightly, relative to White students (for example young people from Black Caribbean and Pakistani families) suggesting that their aspirations/expectations, attitudes and/or behaviours may be slightly less conducive to attainment compared to White students. Conversely, Indian students slightly improve relative to White students (suggesting their aspirations/expectations, attitudes and/or behaviours may be more

conducive). It's worth noting that these changes are very small, however, and worth re-examining with larger samples of ethnic minority groups.

Ultimately, when we include all of the domains in the model, shown by the bottom, purple bar, we would hope to see the effect of ethnicity disappear (i.e. all differences would be explained by intermediary measures such as social and economic position, home environment, attitudes or behaviours, and ideally by measures which could be addressed by policy and advocacy over time). In practice, young people from a Black African and Black Caribbean background did significantly worse than White students (-11.0 and -12.4 points respectively), all other factors held equal. We cannot ascertain the causes of these differences from our models, however three possibilities have been discussed more comprehensively elsewhere[81] and relate to teacher expectations and cultural explanations: Teachers may (consciously or unconsciously) have lower expectations for particular ethnic groups; young people may be disadvantaged by the adoption through peer pressure of an 'urban' or 'street' subculture that is less conducive to getting good grades (particularly by some Black Caribbean pupils); or, according to stereotype-threat theory, by the unconscious adoption of negative group stereotypes which reduce attainment[82]. It's also quite possible that some other disadvantage, that we have been unable to measure here, explains these differences (the model explains 57% of the variation in attainment at the individual level so we know that it does not provide a complete understanding). Either way, the difference in attainment cannot be fully explained by our model, and merits further investigation. The issue of **progress** in attainment throughout secondary school is discussed in Chapter 4 and provides further context to these differences (with indications that these attainment deficits stem, at least in part, from the pre-secondary phase of life).

The above example shows that there are very complex patterns at play and that careful interpretation is necessary. Indeed, it is important to consider these findings within the context of other analyses including, for example, the synthesis report, Ethnicity, Gender and Social Mobility produced for the Social Mobility Commission in 2016, which itself identifies a "Black penalty" in secondary school[83]. It is important not to over-simplify the narrative that seeks to explain differences in attainment across different ethnic groups. Crucially, much of the difference in attainment can be attributed to differences in opportunities (for example coming from a more or less advantaged background). The other domains also correspond with attainment, with each domain being associated with the attainment of young people from differing backgrounds to varying degrees, adding further nuance to the ways in which the importance of ethnicity should be understood.

Home environment

We now consider four aspects of the home environment, whose basic relationship with attainment was summarised in Chapter 2: parental aspirations and expectations; the material support that the home environment provides; parental engagement in their child's education; and quality of the parent-child relationship.

Figure 28 Measures of the home environment associated with attainment
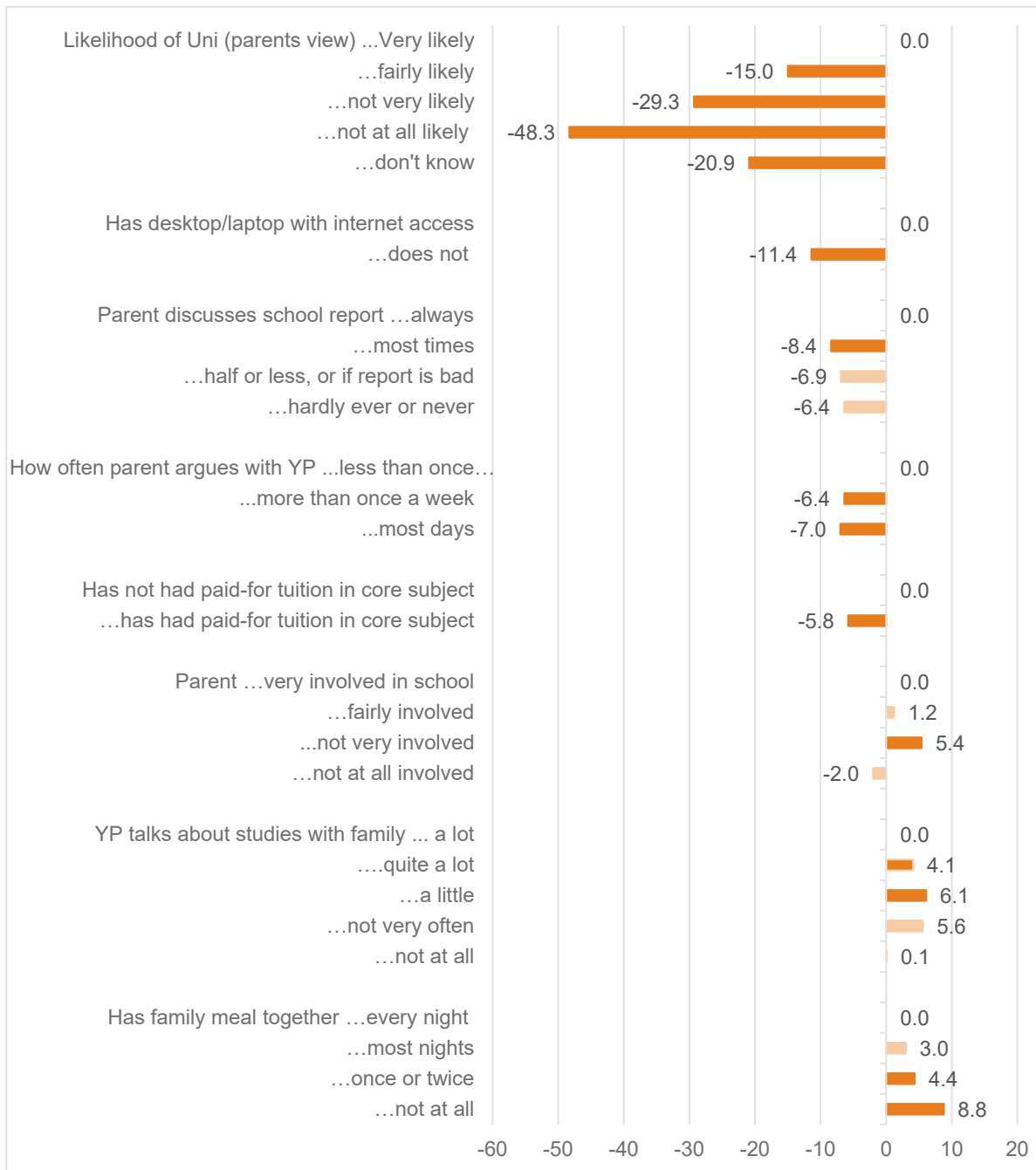
Note: dark bars are statistically significant at a 95% degree of confidence

In terms of **aspirations and expectations**, as shown in Figure 28 above, a parent's view about how likely it was that their child would go to university was highly predictive of the

child's subsequent attainment. As mentioned earlier, this measure was drawn from year 10 data – as such it would pre-date the point at which parents would have known KS4 mock examination or full examination results. Compared to young people whose parents thought they were very likely to go university, those whose parents thought they were not at all likely to go (-48.3 points), not very likely (-29.3), fairly likely (-15.0) or did not know (-20.9) all achieved lower scores on average. At first sight this feels like a tautology; the parent may simply be reporting on how well the young person is doing having already completed several years of secondary education and accurately predicting their achievements ahead of time. In fact, as we will show in Chapter 4, parental expectations regarding university attendance remain predictive of attainment once we take account of their prior attainment at key stage 2, showing that parental expectation is, in itself, a helpful measure.

The second aspect we consider is the **material support that the home environment provides** for the child's education. Figure 28 above also shows that having access to an internet-connected desktop or laptop computer was a key factor in young people's attainment (those who did not have such access attained an average of 11.4 points less than those who did). However, having an internet-connected tablet or smartphone was not found to be significant when all other factors were taken into account (in contrast to the bi-variate findings in Chapter 2). Young people who had received extra tuition in core subjects were seen to have higher attainment in our bivariate analysis, but actually had lower attainment in the final model. What we cannot tell from this particular analysis is whether some of these pupils were receiving tuition specifically because they were struggling in subjects and whether they would therefore have had lower modelled scores had they not received extra tuition. Interestingly, tutoring in non-core subjects was positively related to attainment (+10.2) when first introduced into the model (i.e. before all domains had been included) but ceased to be significant when we took account of the young person's attitudes and behaviours.

The third aspect of the home environment relates to **parental engagement in the child's education**, including: how regularly they discuss school reports; whether they attended parents evenings; whether they felt involved in school life; whether they helped with homework; and whether they talked with the young person about their future study or the day's events at school.

The most important measure in this respect was whether the parent always discuss their child's school report with their child. Young people whose parents discussed their school reports most of the time attained 8.4 fewer points than those whose parents always discussed them. It also appears likely that those whose parents only discussed the reports half the time/if the report was bad/hardly ever/never were similarly disadvantaged compared to those whose parents always discussed their reports (although the differences with these categories were not statistically significant, in part due to the

smaller base sizes of these groups). The significance of whether a parent is involved with school or talks with their child about their future study is considered further below[84].

The fourth and final aspect of the home environment we considered was the **quality of the parent-child relationship** in terms of family cohesion (whether they have family meals together) and parental closeness (whether they talk about things that matter and how often they argued). Here we found that young people did worse if their parents reported arguing with them on most days (-7.0) or more than once a week (-6.4 points) compared to those who argued less frequently.

We also observed some unexpected findings here. Family cohesion was measured by the frequency of eating meals together and contrary to expectations we observed that pupils who only ate meals with their family once or twice a week, or less often, had higher attainment on average (+4.4 and +8.8 points respectively) than those who ate together every day. This may be because households who always eat together have particular characteristics which we have not captured in our other measures, for example, or the quality of interactions during these meals. Similarly, when we looked at self-perceived parental engagement in education we observed that young people did better if their parents reported not feeling very involved in their school life (+5.4 points) or young people reported talking about future studies with family members a little (+6.1 points) or quite a lot (+4.1 points) compared to those who talked a lot. It should be noted that the 'involvement' metric is a measure of the parents' perception of their involvement rather than an objective measure - some of those parents who are actively involved in school may also be more inclined (or able) to identify areas where they could do more, and therefore self-assess their involvement more harshly than others who are, in reality, less involved.

Young person's attitudes and behaviours

As explained in Chapter 2, we took account of a very large number of measures of young people's attitudes and behaviours, which we have categorised into four sub-domains: school behaviours; aspirations and expectations; and risky and prosocial behaviours.

School behaviours were the strongest behavioural predictors of low attainment (see Figure 29). The strongest predictor was truanting, though caution is needed here because only 53 young people in our sample reported truanting more than a few specific lessons. The confidence intervals around these figures are consequently very large, but the much lower levels of attainment amongst this minority group is nevertheless striking; compared to those who did not truant, the apparent 'attainment penalty' for truanting for weeks at a time was -104.5 points, for days at a time was -61.5 points, and for those who did not know or refused to report how much they truanted was -39.0 points. The effect of

---

[84] Help with homework was not significant when included in the full model.

skipping particular lessons was not, however, statistically significant[85]. The other, very much more common non-compliant behaviour in school related to failure to complete set homework. The attainment deficit associated with completing most (rather than all) of the homework set was -10.3 points. Those who only completed some or none of their homework saw a shortfall of -21.7 points compared to those who completed all their homework. Measures of misbehaviour in class and young people's attitudes to school overall were not, however, significant in predicting attainment in our model.

Young people's **aspirations and expectations** were also predictive of their attainment. The sample who intended to work full-time or do something else not in full time education post year 11 was extremely small so we do not quote figures here (n=7). However, their attainment does nevertheless appear to have been markedly lower than that of those planning to continue at school, even having controlled for all the other factors in the model. As might be expected, young people who thought they were not at all likely to apply to university had lower attainment than those very likely to apply (-18.6 points). There was also a significantly lower level of attainment amongst those who were uncertain whether they would apply to university (-21.9 points compared to those who were very likely to apply). There was a similar pattern for those who thought they were not very or not at all likely to get in to university if they applied (-20.4 points compared to those very likely to get in), or who say they don't know (-15.1 points).

It could be argued that reporting likelihood of applying to and getting into university is simply a case of the young person having accurate insight into their likely academic trajectory. On the other hand, both these measures provide significant additional information over and above the other measures which form part of our measurement of the young person's **self-concept**, in particular their belief in their self-ability. Young people who assessed their own school work as below average or not at all good had lower attainment than those who assessed their work as average (-12.3 points). Conversely, young people who thought their school work was very good (+12.4 points) or above average (+7.8) did attain higher scores than those who saw their work as being average. The same applied to those who thought their teacher assessed their work as very good (+14.4 points compared to those assessed as average) or above average (+12.1 compared to those assessed as average). Equating work with success was not found to have a significant relationship with attainment but there was a relationship with locus of control. Young people who had a higher level of locus of control (believing that they have control over the outcome of events in their life) also had higher attainment on

---

[85] Further discussion of the relationship between absence and attainment, based on the full cohort of pupils for whom absence and attainment data were available, are published by the Department for Education: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/509679/The-link-between-absence-and-attainment-at-KS2-and-KS4-2013-to-2014-academic-year.pdf
It should, however, be noted the analyses in the above report are based on a narrower set of variables than was utilised in this report.

average (a one point increase on the 'locus of control' scale, measured on a ten point scale, equated to an increase of 2.4 points in attainment).

Fighting was the only **risky behaviour** that predicted attainment when all other measures were considered (those who had been involved in fighting in the past, without using weapons, scored 9.9 points fewer than those who had not). On the other hand, four **positive or pro-social behaviours** were associated with higher attainment. These were: reading for pleasure (where there was a strong gradient across the scale and those who never read for pleasure attained 14.0 points less than those who read most days); learning a musical instrument (+9.5 points compared to those who did not); playing sport (+8.6 points compared to those who did not); and participating in community work (+6.5 points compared to those who did not). This is a striking finding but one which is open to a number of interpretations. A reasonable hypothesis is that engaging in these extra-curricular activities is beneficial as they may have positive effects on the young person's mood, afford opportunities to socialise, provide additional stimulus and directly enhance certain skills. On the other hand, the direction of causality may not be what it first appears; it is, for example, possible that young people who have more success in school also have more success in activities like sports or music and so are encouraged to continue. At the very least, however, it seems that attainment was positively associated with these four activities. Not all activities saw a significant relationship with attainment. Going to the cinema or a match, playing snooker, darts or pool, going to a youth club and attending religion classes all had a non-significant relationship with attainment after controlling for other factors.

Figure 29 The association of attitudes and behaviours with attainment

SCHOOL BEHAVIOURS

Does not truant
…truants particular lessons — 0 / 4.4
... amount unknown or refused — -39.0
…several days at a time — -61.5
…weeks at a time — -104.5

Does all of the homework set — 0
…most of the homework — -10.3
…some or none of the homework — -21.7

ASPIRATIONS AND EXPECTATIONS

What thinks will do after Year 11... in ft ed — 0
…learn a trade/work-based/apprenticeship — -3.8
…work full-time or something else — -50.4
…don't know, undecided, leave full-time ed — -4.4

Whether YP will ever apply to Uni …very likely — 0
…fairly likely — -5.8
…not very likely — -1.7
…not at all likely — -18.6
…don't know, undecided — -21.9

If applies to university... very likely to get in — 0
…fairly likely — -7.3
…not very or not at all likely — -20.4
…don't know or plans to leave f-t education — -15.1

SELF CONCEPT

Self assessement of school work ... average — 0
…very good — 12.4
...above average — 7.8
…below average or not at all good — -12.3

Thinks teacher assesses work… average — 0
…very good — 14.4
…above average — 12.1
…below average/not at all good — -6.3

Locus of contol — 2.4

RISKY AND PROSOCIAL BEHAVIOURS

Reads on most days — 0
…more than once a week — -4.6
…once a week — -12.0
…less than once a week — -5.4
…hardly ever — -11.2
…never — -14.0

Does not do this activity — 0
…hit/attacked someone (without weapons) — -9.9
…does community work — 6.5
…takes part in sport — 8.6
…plays a musical instrument — 9.5

Note: dark bars are statistically significant at a 95% degree of confidence

85

Health and wellbeing

We also included four measures of health and wellbeing in our models: a subjective report of general health (being very good, fairly good, not very good, not good at all); a measure of psychological distress (GHQ12); a measure of sleep (hours per night); and whether the young person had experienced any of six kinds of bullying in the past 12 months (being called names, being excluded socially, having money or things taken away, being threatened with violence, experiencing violence and experiencing cyber-bullying). Figure 30 shows that three of these measures had a significant relationship with attainment, even when we held constant all other domains in our model.

Figure 30 Measures of wellbeing associated with attainment
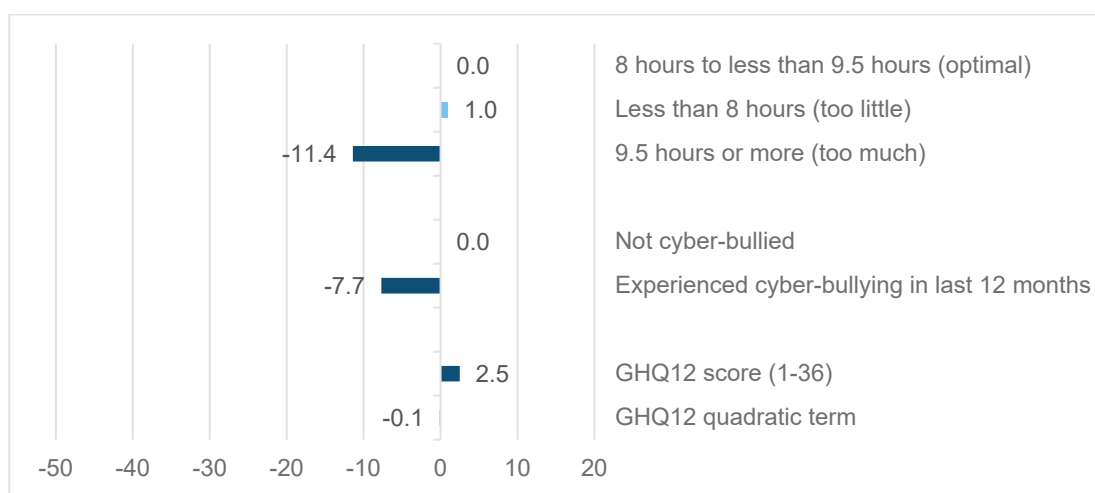
Note: dark bars are statistically significant at a 95% degree of confidence

An interesting finding is that young people who reported sleeping more than the recommended hours of sleep each night (9.5 hours or more) had lower attainment (-11.4 points) than those sleeping less. This may have a causal element where young people who sleep a lot or need to sleep a lot have fewer hours in the day to study, or may be more frequently late for school, for example. However, this is speculative and further research is needed before drawing any conclusions – what is certain is that even after taking account of all the other measures in our model, a sleep effect remains.

It is also striking that, again having controlled for all other measures, experiencing cyber-bullying in the last 12 months accounted for lower attainment (-7.7 points compared to those who had not been cyber-bullied) [86]. None of the other forms of bullying showed a significant effect on attainment after taking account of other factors. The relationship between psychological distress (measured using GHQ12) and attainment is more complex and is described more fully in Box 3 below.

---

[86] It is worth bearing in mind that this measure was drawn from data collected at year 10, so the 'past 12 months' potentially stretches back to year 9.

Box 3 The relationship between psychological distress and attainment is curvilinear

In our previous reporting on LSYPE data we found that young people in this cohort, particularly girls, report significantly higher levels of psychological distress than their counterparts in the first LSYPE cohort nine years earlier. This was supported by other research findings[87] and raises an important question about the short and long term consequences this might have for young people's wellbeing, educational attainment and broader outcomes.

One of our most interesting findings is that in the short term, with all other measures held equal, young people with slightly raised levels of psychological distress appear to achieve higher Best 8 scores. Closer inspection demonstrates that this relationship is curved, as shown below.

Adjusted predictions with 95% confidence intervals
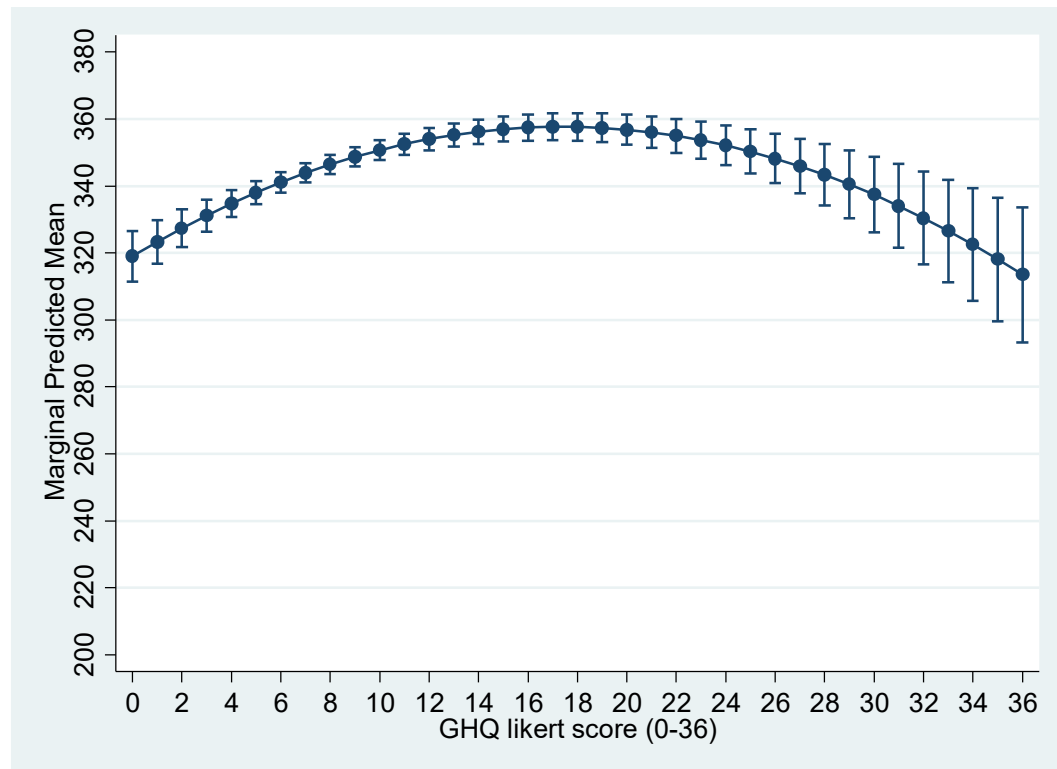
Young people with very low and very high levels of psychological distress go on to achieve lower levels of attainment – and those with moderate levels of distress do better. In popular language we might say that young people who are "too laid back" or "too stressed" are both disadvantaged. Although further research would be needed to clarify or confirm this finding, our evidence is supportive of the hypothesis that "a little bit of stress is a good thing". At the same time it is clear that young people who experience higher levels of psychological distress experience strongly negative outcomes even at this relatively early stage of KS4 examinations.

Addition of school level factors

Alongside the individual level measures described in the preceding sections, we also considered a number of school related factors and found that several had a significant relationship with attainment: school type (whether the school has a selective intake; whether the school has a sixth form); school quality (Best 8 value added); as well as young people's subjective judgements related to their teachers and the attitudes of their peers.

Perhaps unsurprisingly, the strongest predictor amongst these measures was whether the pupil was studying in a non-selective school. This was associated with 20.1 points lower average attainment than was found for pupils in selective schools, even after the other modelled measures were taken into account. It is, however, worth noting that the data in this chapter do not take account of students' pre-existing attainment deficits (i.e. differences arising from the pre-secondary phase of education) as these are highly relevant when considering selective schools – this is discussed further in Chapter 4. The other measure related to school type and which helped to predict attainment was whether the school had a sixth form. An unexpected finding is that young people in schools without a sixth form were predicted to achieve 6.0 points on average higher than those in schools with a sixth form (again, as with all data reported in this chapter, when all other measures were held constant). Although this is a potentially interesting finding it needs to be better understood before drawing any firm conclusions. There may be an explanation based on how schools without sixth forms differ in some other respect, or this may reflect educational policy in different areas or some other measure which has not been included in the model.

We noted in Chapter 2 that there was a bivariate relationship between the proportion of SEN students in a school and the level of attainment. This relationship is not significant when explaining attainment here.

---

[87] For broader discussion of mental health issues in the England see: Mental Health and Wellbeing in England - Adult Psychiatric Morbidity Survey 2014, McManus, Bebbington, Jenkins, Brugha et al, 2016
https://www.gov.uk/government/statistics/adult-psychiatric-morbidity-survey-mental-health-and-wellbeing-england-2014

Figure 31 Measures of school (peers, teachers, structure, quality) associated with attainment
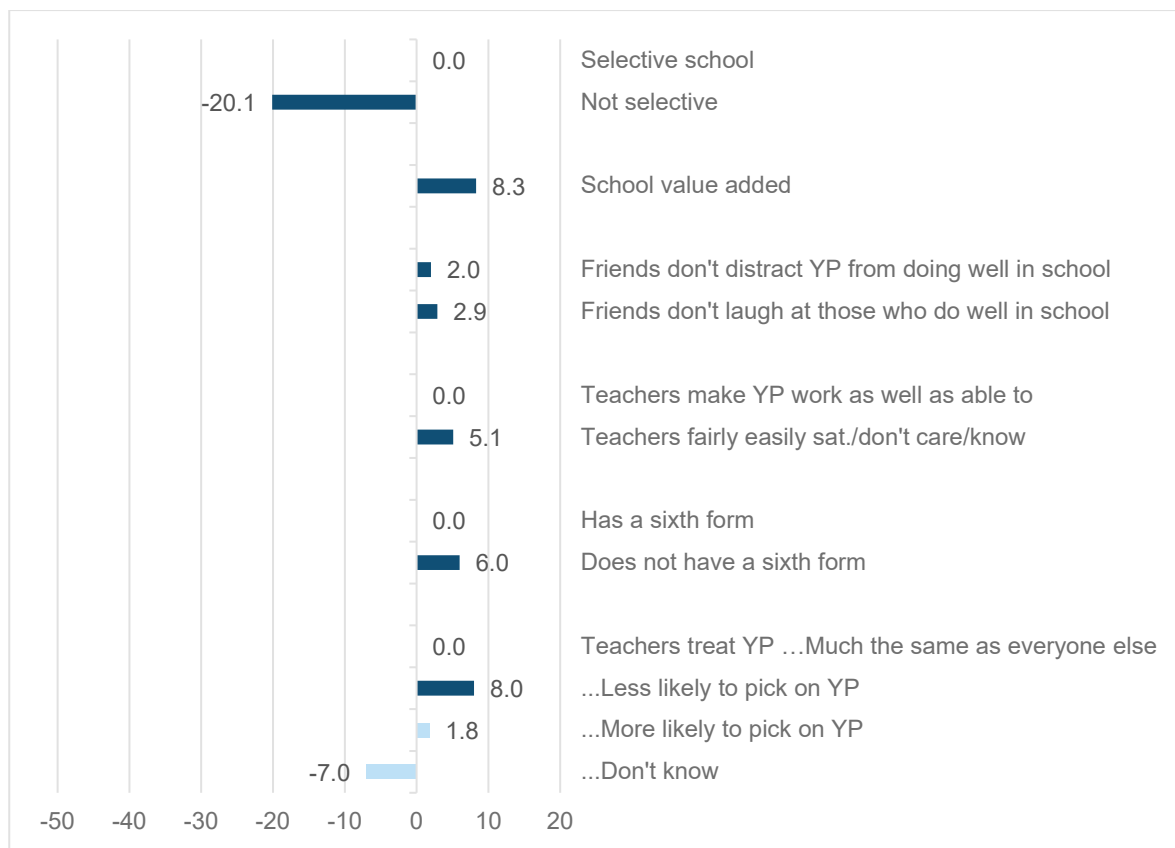
Note: dark bars are statistically significant at a 95% degree of confidence.

Only one measure of school quality was significant when predicting attainment - school added value. An increase in school added value of one standard deviation (which is equivalent to moving up about one-sixth of the population of schools in a league based on this measure) was associated with 8.3 points higher attainment on average, or just over one grade.

In addition to these school level variables, we included student reported perceptions of teachers and peers. Two measures associated with perception of the teachers' behaviour towards the student predicted higher attainment. The first was disagreement with the statement 'Most of my teachers try hard to make me work as well as I am able' (+5.1 points compared to those who agree) which at first sight appears contradictory but may reflect that, when all other factors such as home environment and attitudes and behaviours are held equal, pupils who do not agree with the statement may be reporting that they are more self-motivated. This does not mean that teachers should not encourage students to work hard – when we look at a simple bivariate relationship (Appendix B) we see that students do better on average when they agree with the statement that the teacher makes them work hard; but when we take account of many other factors, we see this pattern reversed. The second measure associated with teachers that predicted higher attainment was that when asked how likely the teacher

was to pick on them, relative to other students, those who said that they were less likely to be picked on attained higher scores (+8.0 points compared to those who felt they were picked on by their teacher much the same as everybody else). In this instance, the positive effect on attainment of being picked on less than others is also observed in the bivariate analysis; being picked on more had a highly negative effect but this was not significant in the model once all other factors had been taken into account. Again, this is open to different interpretations.

Two of the four measures of peer relationships (inevitably, based on subjective reports given by the young people) were also predictive of higher attainment, even when all other measures were taken into account. Looking first at young people's responses to the statement 'friends distract me from doing well in school', for every step on the scale from strongly agree to strongly disagree, attainment increased by an average of 2.0 points. The same pattern is observed in relation to the statement 'friends laugh at those who do well in school'. For every step on the scale from strongly agree to strongly disagree, attainment was predicted to increase by 2.9 points on average.

Addition of measures related to area

We considered including a wide variety of individual level measures relating to the area where the young person lived but ultimately focused on region, IDACI and whether the young person lived in an area that was urban or rural[88], with the significant predictors shown in Figure 32 below.

---

[88] Demark, S., Platts-Fowler, D., Robinson, D., Stevens, A., and Wilson, I. (2010) Young People and Community Cohesion: Analysis from the Longitudinal Study of Young People in England (LSYPE) DfE Research Report No. RR033; Green, R. (2010) Educational Attainment Among Young People in Rural England – A Multilevel Analysis, ICLS Occasional Paper 4.2

Figure 32 Measures relating to the young person's area associated with attainment
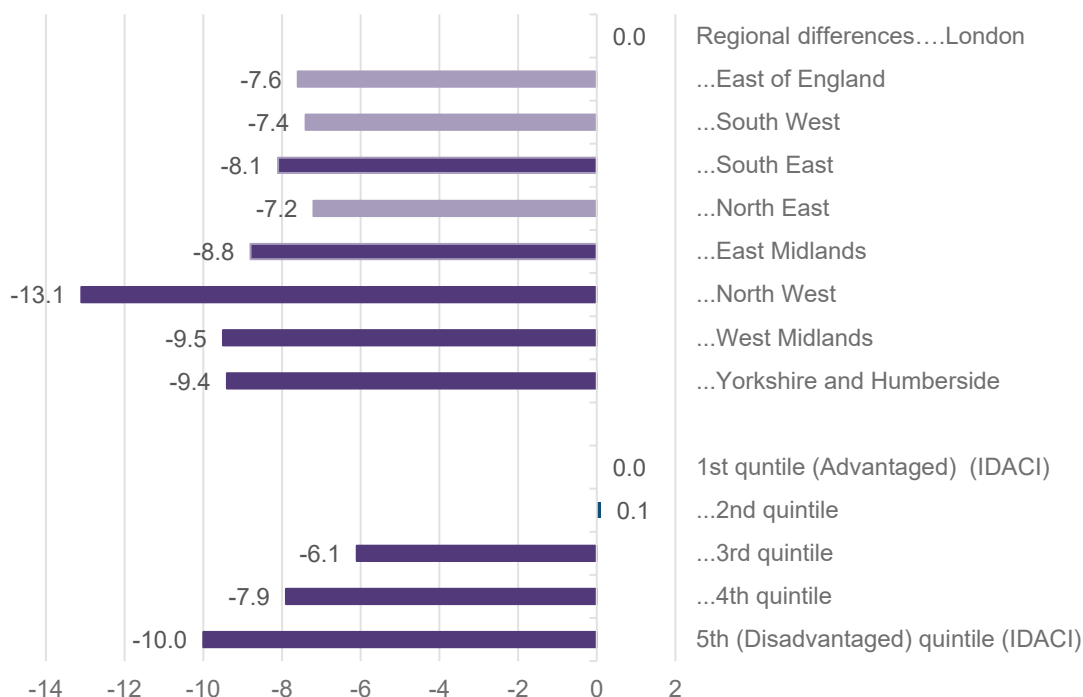
Note: dark bars are statistically significant at a 95% degree of confidence

Young people living in all regions achieved lower scores on average than those in London and we showed in Chapter 2 that the difference was significant in five regions. Our model allows us to control for many underlying differences between these regions, for example the socio-economic circumstances of the students or the student intake in the schools. Even having taken these other measures into account, four of the five areas continued to fare worse than London in terms of attainment, though the difference was less pronounced than that described in Chapter 2. Compared to London there was significantly lower attainment in the North West (-13.1), the West Midlands (-9.5), Yorkshire and Humberside (-9.4) and the East Midlands (-8.8). In addition, the South East had attainment that was significantly lower than London once all other factors had been taken into account (-8.1 points).

Of course, regions are not homogeneous so it is unsurprising that the level of disadvantage in the local area where the young person lives is also predictive of their attainment at KS4. On average, young people in the three most disadvantaged quintiles, as measured by IDACI, were predicted to attain 6.1, 7.9 and 10.0 points fewer (respectively) than those in the most advantaged quintile.

Predicting the Level 2 English and maths threshold

Our main focus in Chapter 3 has been on predicting overall attainment measured on a continuous points scale (the Best 8 scale), which is the total score for each pupil based

on their best 8 GCSE or equivalent results. Here we turn to the achievement of the Level 2 English and maths threshold (whether or not the young person achieved A*-C in both GCSE English and Maths or equivalent) to identify whether the same factors predict this measure of attainment.

Overall, we found that the pattern of results when modelling the Level 2 English and maths threshold was very similar to the findings based on the Best 8 measure. However, in some instances measures which were associated with a difference in attainment using the Best 8 scale did not show any significant difference in terms of the Level 2 English and maths threshold. In part, this is due to the nature of the threshold outcome measure which is binary and so has less statistical power for identifying these differences.

Nevertheless, using the Level 2 English and maths threshold, the measures shown in

Figure 33 were found to be predictive. We present these findings in terms of the increased or reduced odds of attaining the threshold in comparison with the reference category[89].

Personal characteristics

In terms of **personal characteristics,** we showed earlier in this chapter that Black African and Black Caribbean achieved lower Best 8 scores than White students. However, there were no significant differences in the odds of students from these groups reaching the Level 2 English and maths threshold compared to White students. The only statistically significant difference compared to White pupils related to young people with Mixed heritage, whose odds of achieving the Level 2 threshold were 1.8 times lower than those of White pupils (a difference which was not present in terms of the Best 8 measure).

Reinforcing the importance of SEN status shown earlier, the odds of young people with SEN (including statements and/or school action) reaching the Level 2 English and maths threshold were 3.7 times lower than those without SEN. However, disability status was not a measure which helped predict this measure of attainment.

---

[89] To allow for the simple and direct comparison of effect size between positively and negatively associated factors with attainment, we have converted odds ratios for negatively associated factors (which would normally be less than one) so that they are on the same scale as positively associated factors (e.g. an odds ratio of 0.5 would convert to -2.0). We then interpret these in the same way that you would interpret positive odds ratios, that is as a multiplication of the odds.

Figure 33 Predicting the Level 2 English and maths threshold: personal, family and home circumstances (odds relative to reference category)
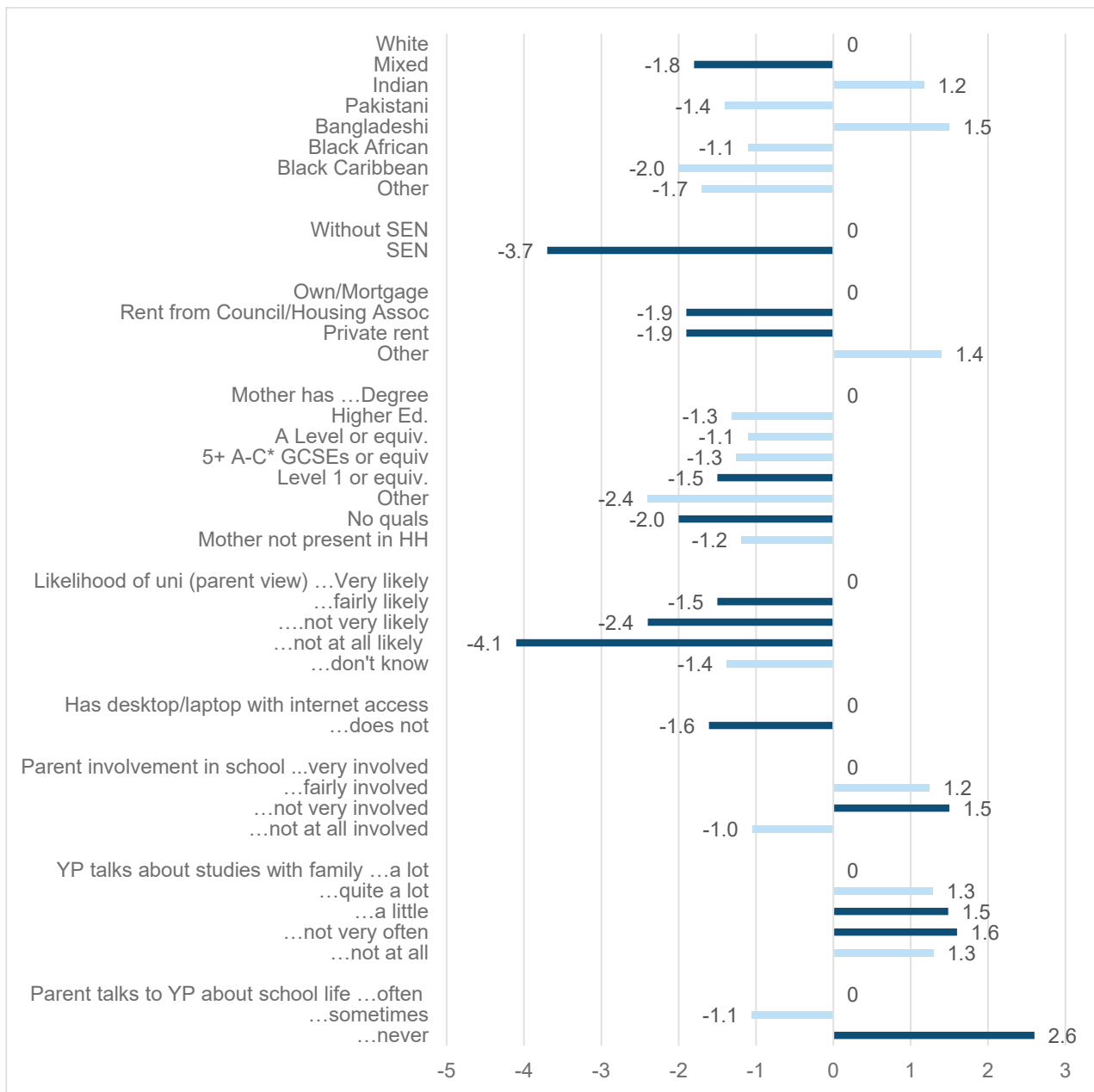
Note: dark bars are statistically significant at a 95% degree of confidence

Family background

Turning to **family background**, the odds of young people who lived in rented accommodation achieving the Level 2 English and maths threshold (regardless of whether they lived in Council or Housing Association, or in the private rental sector) were 1.9 times lower than young people who lived in a home that was owned or had a mortgage. This is very similar to our finding related to the Best 8 outcome measure.

Mother's education was found to be important for predicting attainment based on Best 8 scores and was also important when predicting attainment of the Level 2 English and

maths threshold, albeit with far fewer significant differences. The odds of young people whose mothers had Level 1 qualifications achieving the Level 2 English and maths threshold were 1.5 times lower than for those whose mother held a degree, and were 2.0 times lower for those who had no qualifications than for those whose mother held a degree. Other differences were not statistically significant.

Still focusing on family background, the measures relating to household income and family composition, which had shown significant differences in terms of Best 8 scores, were not significant in terms of the Level 2 English and maths threshold.

Home environment

With regard to **home environment**, the likelihood of young people achieving the Level 2 English and maths threshold varied strongly by the extent to which their parent thought them likely to go to university. Compared to those who were very likely to go to university, the odds of achieving the threshold were 1.5 times lower for those who were thought fairly likely to go to university by their parents, 2.4 times lower for those who were thought not very likely and 4.1 times lower for those who were thought not at all likely to go to university. Young people's odds of achieving the threshold were also 1.6 times lower if they did not have an internet-connected desktop or laptop (than those who did). Consistent with our somewhat unexpected finding earlier, the odds of young people whose parents were not very involved in school life achieving the threshold were 1.5 times greater than those whose parents were very involved. Similarly, in comparison with young people who talked with their family about future studies a lot, the odds of achieving the Level 2 English and maths threshold were 1.5 times greater for those who talked about future studies 'only a little' and 1.6 times greater for those who talked about future studies 'not very often'. Indeed, the odds of those who said they 'never' talked about school reaching the threshold were 2.6 times greater than those who talked about it often, all other factors held equal.

Some home environment factors that helped predict attainment of Best 8 GCSE or equivalent scores were not found to predict attainment of the Level 2 English and maths threshold. These included how often the parent read school reports, how much they argued with the young person and whether they paid for core tuition.

Attitudes and behaviours

Turning to the young person's own **attitudes and behaviours** (see Figure 34), the odds of young people achieving the Level 2 English and maths threshold were 1.5 times greater if they thought their teacher assessed their work as very good and 1.6 times greater if they thought the teacher assessed their work as above average (compared to those who thought the teacher judged their work to be average). The odds of achieving this threshold was also greater among those who had a higher locus of control (a one unit increase on a ten point scale used to capture this belief increased the odds of

achievement by 1.2 times). Relative to those who were very likely to apply to university, the odds of attaining the Level 2 English and maths threshold were 1.4 times lower for those who were only 'fairly likely' to apply to university and were 2.9 times lower for those were didn't know or were uncertain. Being hesitant about whether they would get into university *if they applied* was also associated with lower odds of attaining the threshold - the odds for those who thought that they were not very likely or not at all likely to get in if they applied were 2.4 times lower than for those who thought they were very likely to.

Some elements of school behaviour were very important when predicting Best 8 attainment but truanting was not significant in terms of the Level 2 English and maths threshold (or at least we did not have the statistical power to be able to detect a difference). There is, however, some evidence that homework remained an important predictor; the odds of young people achieving the threshold who reported only doing most of their homework were 1.4 times lower than for those who reported doing all of it, and 2.3 times lower among those who reported doing some or none of it.

As was the case in our analysis of Best 8 scores, those who read less regularly than daily had lower attainment in terms of the Level 2 English and maths threshold. For example, the odds of achieving the threshold for those who never read for leisure were 1.5 times lower than for those who read daily and the odds for those who read for leisure about once a week were 1.5 times lower, also in comparison with daily readers.

As might be expected based on the findings presented earlier, the odds of those involved in fighting achieving the Level 2 English and maths threshold were 1.8 times lower than those who did not. At the same time, the prosocial activities mentioned as important in achieving higher Best 8 scores were also important as predictors of attaining the Level 2 English and maths threshold. The odds of doing so were 1.9 times higher for young people who played a musical instrument than those who did not, 1.5 times higher if they did community work compared to those who did not and 1.4 times higher if they played sport compared to those who did not. In addition, the odds of achieving the threshold were 1.3 times higher if young people went to the cinema, theatre or concerts (these had not been significant activities when predicting attainment in terms of Best 8).

Figure 34 Predicting the Level 2 English and maths threshold: attitudes and behaviours, wellbeing, school and area
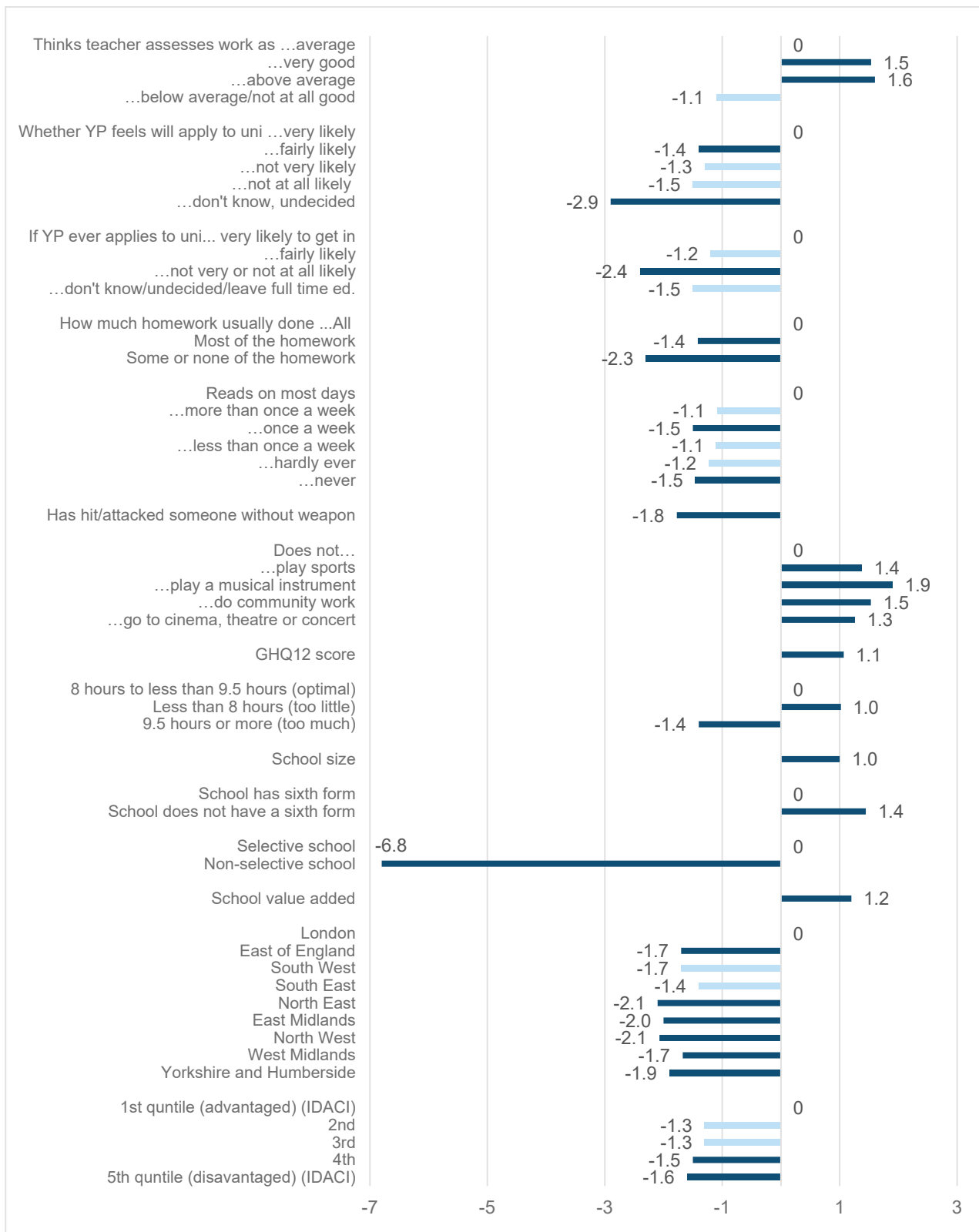
Note: dark bars are statistically significant at a 95% degree of confidence

Health and wellbeing

When looking at **health and wellbeing**, two measures helped to predict attainment of the Level 2 English and maths threshold. First, the odds of young people who reported sleeping 9.5 hours or more achieving the threshold were 1.4 times lower than those who reported sleeping the recommended amount (between 8 and less than 9.5 hours). Second, at first sight it appears that young people with higher levels of psychological distress as measured by GHQ12 were also more likely to achieve the Level 2 English and maths threshold. In Box 4 in Chapter 3 we showed that the relationship between psychological distress and the Best 8 measure of attainment was 'curvilinear', with poorer attainment being associated with both the lowest and highest levels of psychological distress. In this case the squared term, which indicates whether the relationship is curved, is not significant, and suggests the odds of reaching the Level 2 English and maths threshold is increased 1.1 times for each 1 point increase in GHQ score. Unlike the predictions of attainment measured by Best 8 scores, bullying was not found to be a significant factor in predicting attainment of the Level 2 English and maths threshold.

School measures

Turning to **school** measures, the odds of young people in non-selective schools achieving the Level 2 English and maths threshold were 6.8 times lower than those in selective schools. This very substantial difference probably reflects the fact that young people who are likely to fail Maths and English GCSE or equivalent are unlikely to get into selective schools or unlikely to remain in those schools if they are not maintaining good academic performance. The odds of those in a school with no sixth form achieving the threshold, again controlling for all other factors, were 1.4 times greater than those in a school with a sixth form. An increase in school added value of one standard deviation (which is equivalent to moving up about one-sixth of the population of schools in a league based on this measure) was associated with the odds of achieving the Level 2 English and maths threshold increasing by 1.2 times.

Area measures

Young people in most **areas** were less likely than those in London to reach the Level 2 English and maths threshold, significantly so in North East (the odds were 2.1 times lower than those in London), North West (2.1 times lower), East Midlands (the odds were 2.0 times lower), Yorkshire and Humberside (1.9 times lower), East of England (1.8 times lower) and West Midlands (1.7 times lower). The level of disadvantage in the area, based on IDACI, was also found to be predicative of Level 2 English and maths attainment with the odds of young people in the fourth quintile achieving the threshold 1.5 times lower, and in the fifth quintile 1.6 times lower, than those in the most advantaged areas.

Chapter 4: Explaining progress from KS2 to KS4

In this chapter, instead of considering attainment per se, we identify the measures that predict *school progress* in terms of the Best 8 scale. This is the relative progress pupils make during secondary school, measured in terms of GCSE (or equivalent) grades.

We achieve this by introducing an additional measure to our statistical model - the pupil's attainment at KS2 based on the English and Maths Standard Attainment Tests (SATs) administered in year 6[90]. By including KS2 as an additional measure we 'control' for prior attainment, which is both a marker of where the child sits academically, and of prior advantage (as children from more advantaged backgrounds tend to benefit more from the schooling that they are provided with). In other words, the 'control' adjusts for differences in young people's attainment at the point they begin their secondary education.

When it comes to interpreting the data in this chapter, measures that are significant *after* we control for KS2 can be understood as predictors of the progress a child makes *in the time between the KS2 tests and KS4 examinations*. This indicates the progress a young person makes as a result of what happens within secondary school itself but will, of course, also reflect the broader set of advantages and opportunities they may experience outside of school during this period.

We 'interpret' progress in terms of the same points scale presented in Chapters 2 and 3, which can be converted to GCSE grades. As such, when we refer to 'progress', we mean the extent to which the Best 8 score associated with a particular characteristic is higher or lower than a relevant comparator (e.g. boys vs. girls) after controlling for differences in their relative performance at KS2.

Before continuing, we need to note that key stage 2 exams were not completed by all young people which has some methodological consequences for the work reported in this chapter. In 2010, two teachers' unions voted for a boycott of the SATs for year 6 pupils because of concerns about too strong a focus on testing and disquiet about school league tables and their impact on schools with disadvantaged pupils. A substantial number of schools took part and, as a result, the LSYPE2 data set has missing KS2 data for approximately 30% of the cohort. Given the necessity of having KS2 scores for measuring progress through the school system, DfE funded RAND Europe and Cambridge University to construct imputed data sets which could be used to approximate

---

[90] The meaning and significance of KS2 scores have been widely debated and the value of the measure has been disputed by some educationalists. Arguably KS2 scores capture a combination of a number of phenomena: a child's aptitude and drive; the quality of the primary school they attended; the focus of that school in preparing for the SATs; and, as has been shown in prior research, differences in the social and economic circumstances, home and local environment, attitudes, behaviours and wellbeing of the child and their parents, as well as the characteristics of the rest of the school's pupil intake which affects the demands placed on that school and the quality of education it is able to provide.

KS2 attainment in projects such as ours. A carefully developed and valid method of data imputation was employed that adequately accounts for missingness (see Saunders et al, 2016 for further detail of the methods employed) [91].

Consequently, we linked imputed KS2 data to our survey data sets for the purpose of this analysis. Since the recommended approach to using this data is to incorporate multiple variants of the imputed data, we employed a specific kind of regression model which could handle this particular complexity. Further detail is provided in Appendix A Methodology.

Prior attainment at KS2

When KS2 is introduced into the model it is, unsurprisingly, a strong predictor of attainment. For every one point increase in KS2 score we see an 8.9 point increase in Best 8 scores, all other factors held equal[92]. Once KS2 has been introduced into the model, the association of the measures with attainment discussed throughout Chapter 3 is changed. We describe this in greater detail throughout the remaining sections of this chapter.

Personal characteristics and family background

Figure 35 below shows the personal characteristics which serve as significant predictors of progress (dark bars) and those which are not significant (light bars). Previously, as described above, we found that boys had lower attainment overall than girls. What we also find is that boys made significantly less progress (-7.5) during secondary school compared to girls. In other words the attainment gap between girls and boys appeared to increase during secondary school[93].

---

[91] A technical report for this work can be found here:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/578541/20161205_Technical_report_FINAL.pdf

[92] Previous research has also identified a strong relationship between attainment at KS2 and KS4. See: Analysis of use of Key Stage 2 data in GCSE predictions, Benton and Sutch, 2014
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/429074/2014-06-16-analysis-of-use-of-key-stage-2-data-in-gcse-predictions.pdf

[93] The pattern underlying this may be complex, as suggested by other research into this topic. For example, work by the Cambridge Assessment Research Division has identified variations in progress by subject from KS2 to KS4, with girls catching up with boys in maths, overtaking them in science and essentially maintaining their higher attainment in English. See: The changing gender gap, Cambridge Assessment Research Division, 2015
http://www.cambridgeassessment.org.uk/Images/the-changing-gender-gap-graph.pdf

Figure 35 Explaining progress: personal characteristics
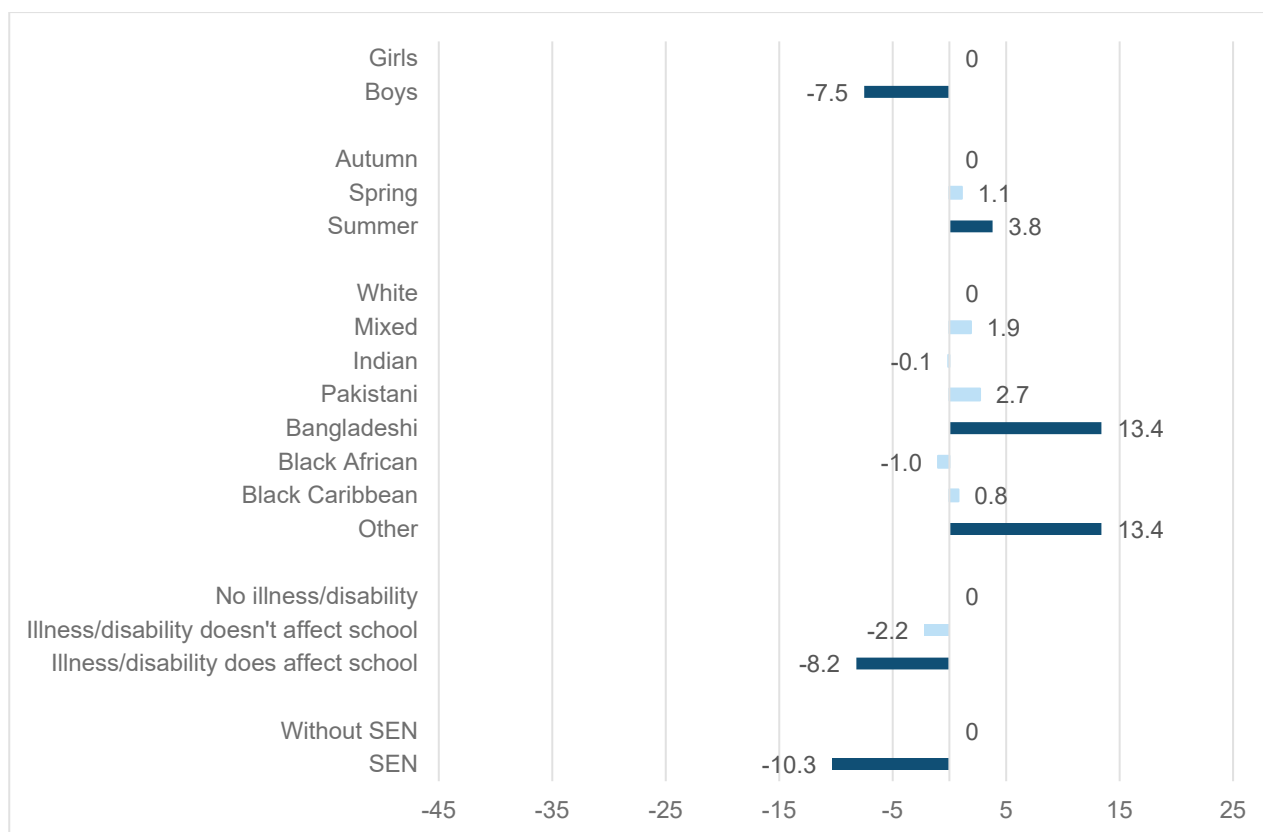
Note: dark bars are statistically significant at a 95% degree of confidence

It is also interesting to note that young people born in the summer term made significantly *more* progress in the five years of secondary school (+3.8) than students born in the Autumn term – the younger pupils were, effectively, catching up during this period.

Earlier, in Chapter 3, we showed that having SEN status was associated with having significantly lower attainment (-34.7 points less on average than those without SEN). Once we take account of prior attainment at KS2 we see that SEN pupils (including statements and/or school action) also made less progress during secondary school than those who do not have SEN (-10.3 points less on average). Having an illness or disability that affects schooling was also associated with lower progress (-8.2 points compared to those with no illness or disability).

In Box 2 in Chapter 3, we showed that the relationship between ethnicity and attainment is complex and that Black Caribbean and Black African students fared significantly worse than White students once all other factors were taken into account. However, once we adjusted for differences in prior attainment, we found that there was no difference in the amount of school progress made between Black Caribbean, Black African and White students during secondary school (remembering that they had different starting points) – as such, the eventual difference in KS4 performance may stem from earlier life

101

circumstances or experiences (prior to secondary school) [94]. Two BAME groups were found to have made more progress than White pupils: Bangladeshi students (+13.4 points) and Other ethnic minorities (+13.4) [95], once other factors were taken into account [96].

Figure 36 below shows that family background continues to drive differences during secondary school.

- Of all the factors relating to family background, living without either biological parent had the strongest negative association with school progress (-18.1 points, equivalent to approximately 3 grades). It should, however, be noted that the sample size for this sub-category is low (44 in the modelled data), increasing the potential margin of error around the quoted figure.
- Living in the private rented sector was associated with less progress during secondary school than living in an owned or mortgaged home (-9.0 points), as was living in Council or Housing Association rented property (-7.7 points). This mirrors the trend observed with attainment overall.
- Living in a household with a lower than median income, or past or current eligibility for FSM, predicted lower progress. The progress of pupils currently eligible for FSM was 10.8 points lower than that of young people living in households with above median income.
- The progress of young people whose mother had a degree level qualification was typically greater than those whose mothers were less highly qualified (not all differences were significant, though in some instances this may have been as a result of relatively low sample sizes rather than a genuine absence of difference). For example, those whose mother had no qualifications made an average of 9.3 points less progress than those whose mother had a degree level qualification.

---

[94] Steve Strand's work on the first LSYPE cohort has previously noted a similar finding that Black Caribbean pupils start behind White British pupils at KS2 and make the same progress throughout secondary school, with the result that they remain behind at KS4.
See: Minority Ethnic Pupils in the Longitudinal Study of Young People in England, Strand, 2008
http://dera.ioe.ac.uk/7916/1/DCSF-RR029.pdf
School effects and ethnic, gender and socio-economic gaps in educational achievement at age 11
https://www.tandfonline.com/doi/abs/10.1080/03054985.2014.891980
[95] The 'other ethnic minority' group consisted primarily of young people from an East Asian or Arab background.
[96] Descriptive analysis of ethnicity and progress for the 2015/16 cohort is published as an SFR. This shows that young people from Indian and other Asian backgrounds show similar progress to those from a Bangladeshi background in absolute terms (i.e. without controlling for other factors). See:
https://www.gov.uk/government/statistics/revised-gcse-and-equivalent-results-in-england-2015-to-2016

Figure 36 Explaining progress: family background
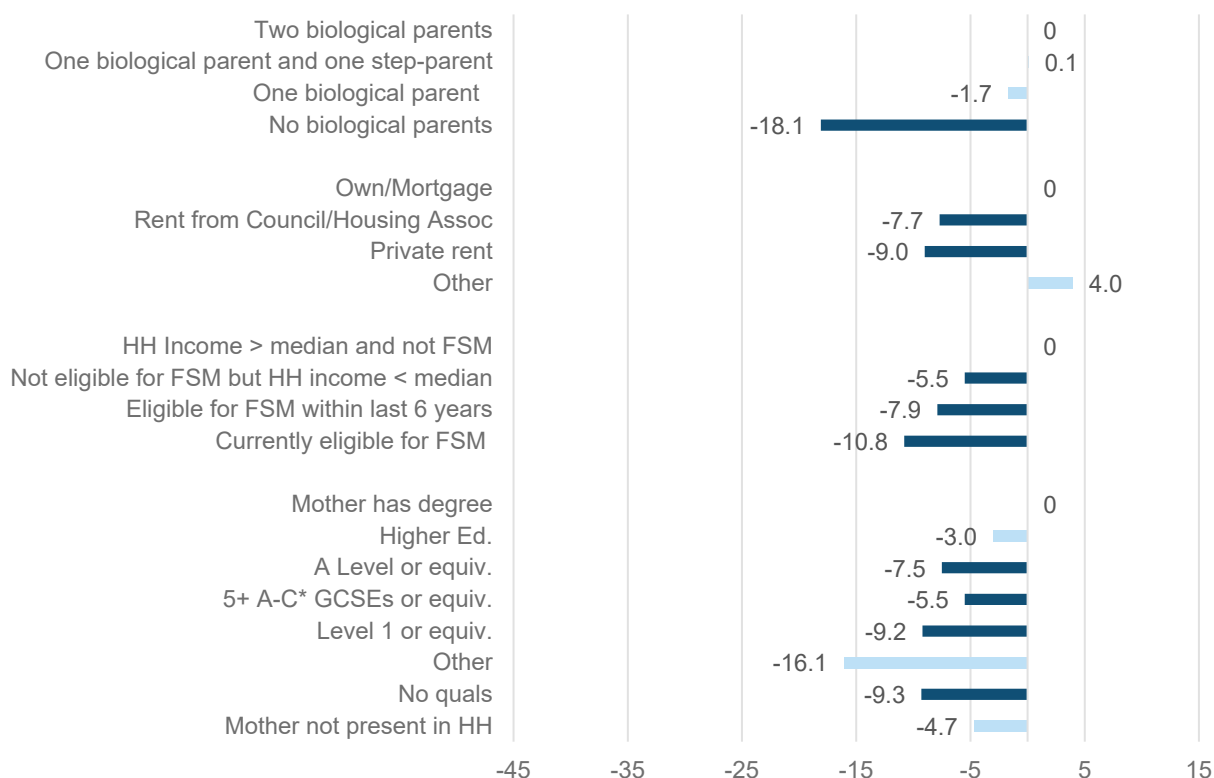
Note: dark bars are statistically significant at a 95% degree of confidence

Home environment, attitudes and behaviours and health and wellbeing

In Chapter 3 we identified eight aspects of the **home environment** which predicted differences in attainment. Four of these were also found to predict lower levels of progress during secondary school: the parent having a negative or uncertain view of whether the young person was likely to attend university (ranging from -12.7 to -27.9 points) [97]; the young person not having access to an internet-connected desktop or laptop (-10.2 points); the parent not always discussing the school report (-6.5 points compared to discussing the report most times); and parent-child arguments (those who argued with their parents more than once a week made less progress than those who argued less often (-4.8 points)).

Furthermore, for reasons not entirely clear, the association between parents *not* feeling very involved in school and higher attainment was also observed when predicting progress (+4.9 points). As we noted in Chapter 3, one possible explanation is that high

---

[97] The main parent was asked their view in Year 10, substantially before the young person sat their KS4 exams, so this measure is not influenced by attainment at GCSE or equivalent. Nevertheless, the very large impact of this measure may reflect, in part, the fact that parents will have a reasonable sense of their child's attainment, and even progress.

levels of parental involvement may occur when a child is doing less well with their school work. It should also be noted that this is a measure of the parents' perception of their involvement rather than an objective measure - some of those parents who are actively involved in school may also be more inclined (or able) to identify areas where they could do more, and therefore self-assess their involvement more harshly than others who are, in reality, less involved.

The three remaining measures which were found to predict overall attainment in Chapter 3 – receiving tuition in a core subject, talking with parents about future studies, and frequency of having a family meal together – were not statistically significantly in predicting progress and so are not shown in Figure 37.

Figure 37 Explaining progress: home environment
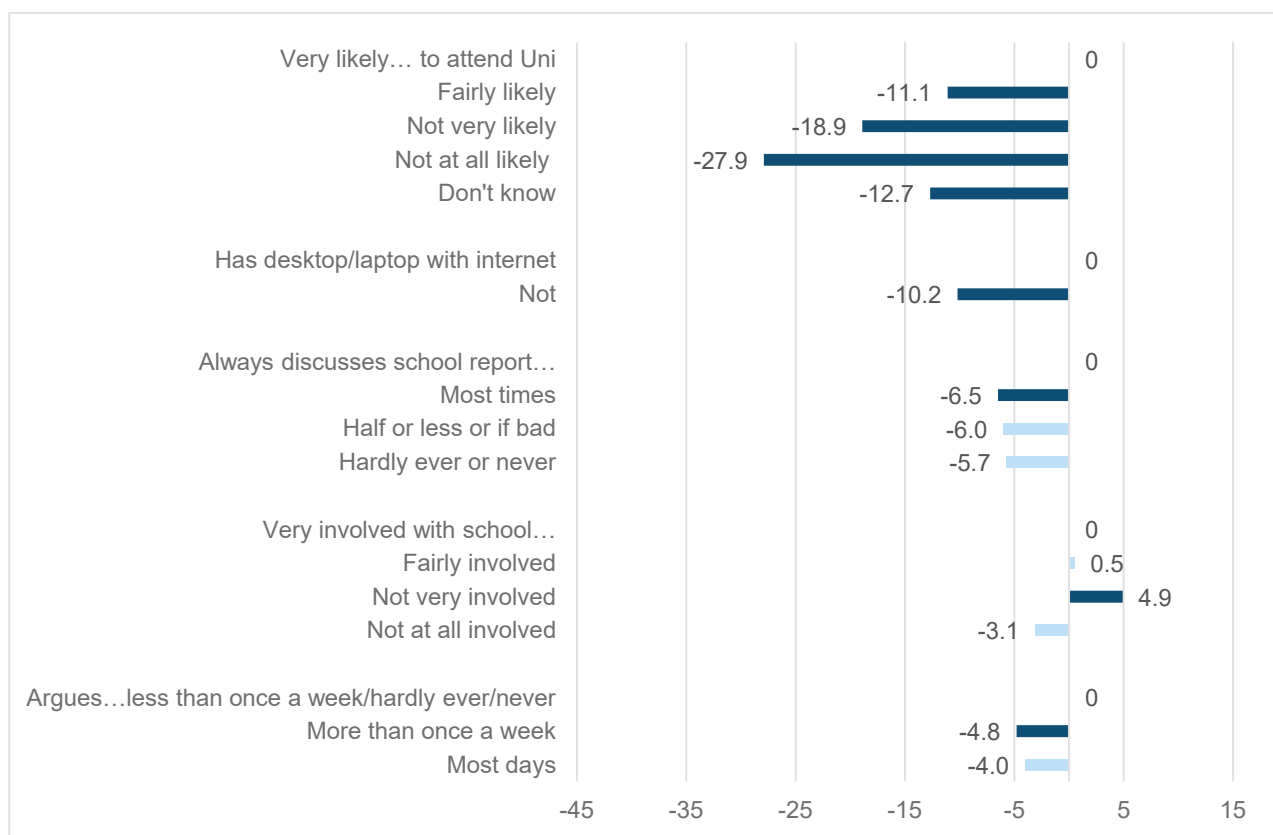
Note: dark bars are statistically significant at a 95% degree of confidence

Turning to the next domain, for the most part the same **attitudes and behaviours** that predicted differences in attainment (see Figure 29) also predicted differences in progress made between KS2 and KS4.

School behaviour was important. In particular, truanting was associated with much lower levels of progress (-76.8 points on average among young people truanting for weeks at a time, -41.4 for those who truanted days at a time and -27.8 who did not know or refused to say how often. However, as explained earlier, the absolute size of these differences should be treated with caution because of the very small numbers exhibiting

such extreme truanting behaviour. Unsurprisingly, not completing all homework also had a negative effect on progress (e.g. those who reported doing some or none of the homework they were set scored 14.5 points less on average, than those who completed all of their homework). Whilst misbehaving in class did not predict attainment it did predict school progress - those who reported misbehaving in half or more classes made less progress (-7.4 points) on average than those who did not.

Figure 38 Explaining progress: attitudes and behaviours

SCHOOL BEHAVIOURS

Does not truant
0
…truants particular lessons
2.1
... amount unknown or refused
-27.8
…several days at a time
-41.4
…weeks at a time
-76.8*

Does all of the homework set
0
…most of the homework
-6.7
…some or none of the homework
-14.5

Does not misbehave in class
0
…misbehaves now and then
-1.4
…in half or more than half of classes
-7.4
…in most or all of classes
-10.5

ASPIRATIONS AND EXPECTATIONS

Very likely to apply to university
0
…fairly likely
-1.7
…not very likely
3.2
…not at all likely
-8.0
…don't know, undecided
-13.1

If applies to university... very likely to get in
0
…fairly likely
-3.6
…not very or not at all likely
-11.4
…don't know or plans to leave f-t education
-9.2

SELF CONCEPT

Pupil assesses work as… average
0
…very good
5.3
…above average
2.6
…below average/not at all good
-3.6

Thinks teacher assesses work as… average
0
…very good
9.7
…above average
7.6
…below average/not at all good
-5.8

Locus of control
-1.7

RISKY AND PROSOCIAL BEHAVIOURS

Reads…most days
0
More than once a week
-3.0
Once a week
-7.5
Less than once a week
-4.2
Hardly ever
-6.3
Never
-6.9

Does not do this activity
0
…does community work
4.5
…plays sport
4.5
…plays a musical instrument
6.1

YP currently smokes
-9.3

Has attacked/hit someone (without weapons)
-8.5

-50   -40   -30   -20   -10   0   10   20

*This bar has been shortened for presentational reasons

106

The predictive power of a young person's **aspirations and expectations** was less strong in terms of predicting school progress than predicting attainment overall. This may suggest that the associated differences in attainment had pre-dated secondary school, although further research would be required to be certain. There was no difference in progress between pupils who wanted to continue in education after year 11 and those wanting to work or do something else. Other factors relating to the young person's self-concept were predictors of progress, although their effect was somewhat weaker. Positive or prosocial behaviours such as reading, playing sport, playing an instrument or doing community work were also found to be less important factors in terms of progress than had been the case for attainment (though they remained significant). Being involved in fighting (that is hitting or attacking someone without a weapon) was, however, as predictive of impaired progress (-8.5 points) as it was of lower attainment. In addition, the young person being a current smoker predicted lower progress (-9.3).

The final domain related to the individual characteristics of the pupil is their **health and wellbeing** (not charted). In Chapter 3, we noted that sleeping longer than our 'optimal' band of 8 to less than 9.5 hours, and experiencing cyber-bullying, were both associated with lower levels of attainment; these factors were not significantly associated with progress. On the other hand, the effect of psychological distress on progress was much the same as its effect on attainment (see Box 3 in Chapter 3), with intermediate levels of distress being associated with most progress and those with high or low levels of psychological distress showing less progress, suggesting that its effect on attainment was evident throughout this period.

School and area effects which predict progress

A number of the measures that relate to **school structure** were found to predict school progress as shown in Figure 39 below.

- Attending a non-selective school predicted a lower level of progress, with those in non-selective schools making 9.2 points less progress than their peers in selective schools[98].
- After controlling for all other factors, pupils in schools with **no sixth form** were found to have made slightly more progress while in secondary school than those attending schools with a sixth form (+4.6 points).
- Interestingly, young people in schools with a **higher proportion of FSM eligible students** progressed slightly <u>more</u> than those in schools with a lower proportion. An absolute increase of 10 percentage points in the FSM eligible students in a

---

[98] This is likely to reflect the nature of these schools which, by definition, select children they predict will be more likely to attain higher results. See, for example, Coe, R., Jones, K., Searle, J., Kokotsaki, D., Mohd Kosnin, A. and Skinner, P. (2008) Evidence on the effects of selective educational systems A report for the Sutton Trust
https://www.suttontrust.com/wp-content/uploads/2008/10/SuttonTrustFullReportFinal11.pdf

school is associated with an increase in the average individual progress by 2.7 points in Best 8 attainment.

Only one measure of school *quality* predicted progress - **school value added**. An increase in school added value of one standard deviation (which is equivalent to moving up about one-sixth of the population of schools in a league based on this measure) was associated with 8.1 points greater progress on average, or just over one grade. The strength of this association is not especially surprising given that school added value is a measure of the relative progress young people make for a given school.

Two measures relating to peers that had been found to predict attainment overall were *not* significant predictors of progress (whether 'friends distract me from doing well at school' and whether 'friends laugh at those who do well in school'). However:

- Two measures of the young person's **relationship with teachers** helped to predict variations in both attainment <u>and</u> progress. Young people who reported being less likely to be picked on by their teachers progressed 5.6 points more on average than those who felt they were picked on much the same as others. Perhaps surprisingly, those who did <u>*not*</u> agree that their teachers "try hard to make them work as well as they are able to" progressed 5.2 points more than those who agreed with this statement. One possible explanation is that young people who were strongly self-motivated experienced less motivating behaviour from their teachers, or were perhaps less aware of it, though other interpretations are possible.
- A third measure of the young person's relationship with teachers, liking none or hardly any of their teachers, also predicted progress (-8.1 points).

Figure 39 Predicting progress: school measures

Measure	Value
School is selective	0
School is not selective	-9.2
School has sixth form	0
No sixth form	4.6
School FSM (Increase of 10 percentage points)	2.7
School Added Value	8.1
Teachers make me work hard	0
Teacher does not make me work hard	5.2
Teacher picks on me…much the same as others	0.0
Less likely to pick on me	5.6
More likely to pick on me	-0.9
Don't know	-2.5
Likes all/most/some teachers	0
Likes none/hardly any teachers	-8.1
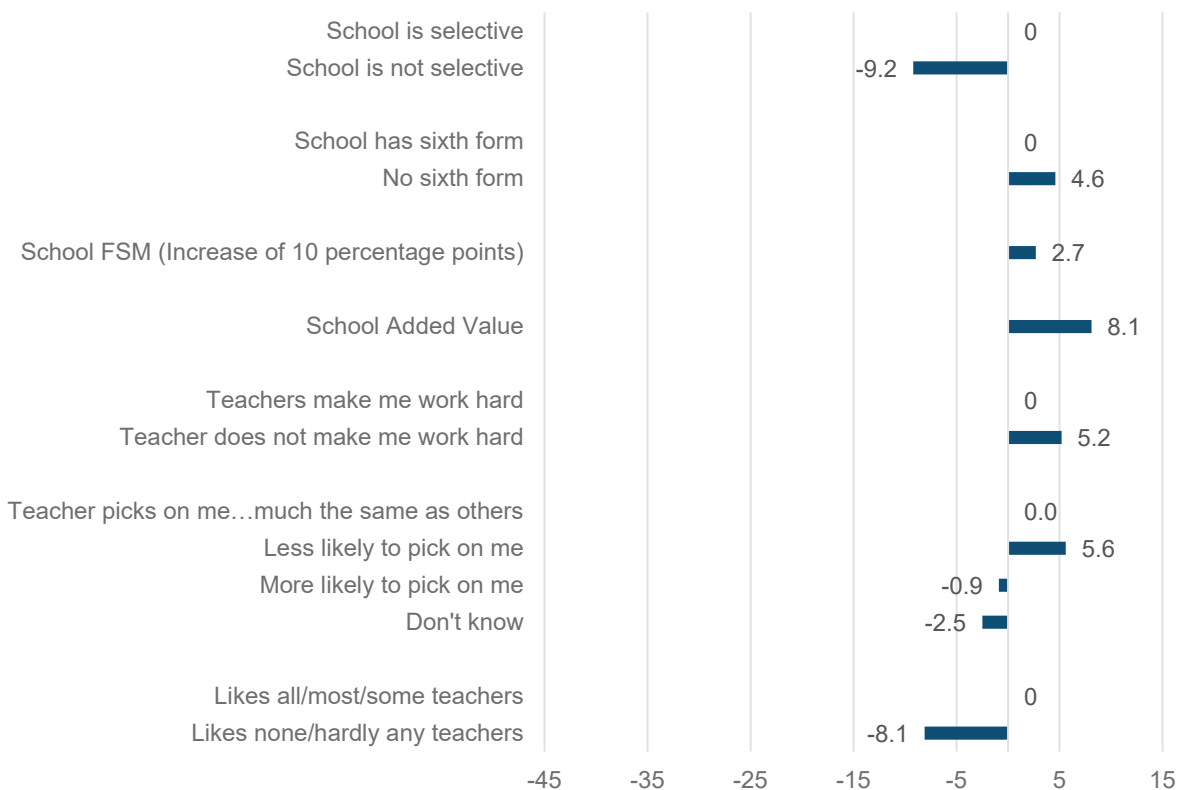
Note: dark bars are statistically significant at a 95% degree of confidence

The effect of living in more disadvantaged areas on progress was slightly less pronounced than the effect on attainment overall, as discussed in Chapter 3. This suggests that living in disadvantaged areas continues to disadvantage young people throughout their schooling period. The measures that predict progress are shown in Figure 40 below. When considering attainment, we found that all areas fared poorly in comparison with London, although only in four areas was the difference in attainment statistically significant. When looking at progress, the association was reduced slightly, with the negative association with living in the East of England and Yorkshire and Humberside no longer statistically significant, but remaining for the West Midlands and North West. Similarly, living in the three most disadvantaged area quintiles, as measured by IDACI, negatively predicted attainment, but only the two most deprived quintiles were significantly associated with progress (-4.7 points for those in the fourth quintile and -6.5 points for those in the most deprived quintile).

Figure 40 Predicting progress: area measures
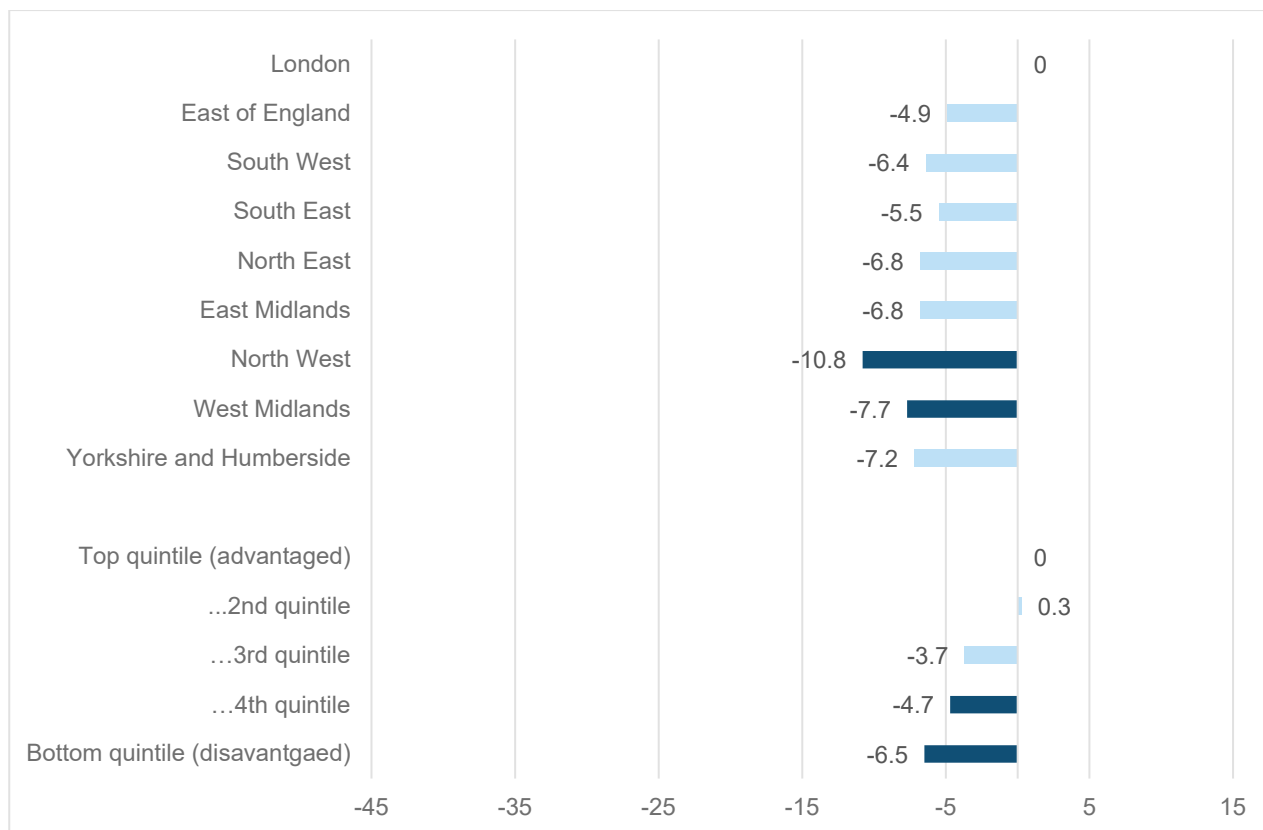
Note: dark bars are statistically significant at a 95% degree of confidence

Summary

In this chapter we have looked specifically at differences in the progress young people make during secondary school, by adjusting for differences in their prior attainment at key stage 2. In practice, many of the factors that explained attainment overall also explained differences in the level of progress young people made during secondary school.

We have drawn attention to a few interesting exceptions that were not evident when we looked at attainment overall. For example, we found that Black Caribbean and Black African students had significantly lower attainment overall than White students, even after adjustment for a broad range of other factors. However, we found no difference in terms of their progress between KS2 and KS4, suggesting that differences in attainment stem primarily from earlier life circumstances or experiences. On the other hand, Bangladeshi and 'Other' ethnic minority groups appeared to progress particularly well (relative to White students) and this would potentially merit further follow-up research.

When considering the school environment, reporting that you do not like any or most of the teachers had a significantly negative relationship with progress (-8.1 points) in spite of our finding that this was not significant when predicting attainment. There are a number of possible interpretations of this finding, but it may support research which

shows the importance of teacher encouragement, and perhaps points to the value of addressing poor relationships[99].

[99] Alcott, B. (2017) Does Teacher Encouragement Influence Students' Educational Progress? A Propensity-Score Matching Analysis Research in Higher Education 58: 773.
https://doi.org/10.1007/s11162-017-9446-2

Chapter 5: Conclusions

The progression of the analysis through the preceding chapters has allowed us to tell an interesting story about the predictors of attainment and progress, drawing upon the richness of the LSYPE dataset. In this final chapter, we summarise the findings that have stood out as important or new and suggest how they can be considered in policy terms.

Complexity beneath face-value differences

The findings which have been outlined throughout this report, while complex, do have a relatively simple theme at their core. Put simply, before placing too much weight on the face-value differences in attainment between different groups of young people, it is important to consider what other factors might be at play.

Our straightforward (bi-variate) analysis in Chapter 2 identified a wide range of performance gaps in the KS4 attainment of different groups when looking at each in turn. In absolute terms, the attainment of those from less advantaged circumstances was lower than that of those from more advantaged circumstances. The attainment of girls was higher than that of boys. Those born in the summer term had slightly lower attainment than those born in the autumn term. Indian and Bangladeshi students had higher average attainment scores than White students who, in turn, had higher attainment scores than those from a Black Caribbean background, and so on.

However, useful though it is to understand the variation in attainment of different groups, our further analysis in Chapter 3 shows that many of these differences are mitigated when other factors are taken into account. For example, further analysis of the gender gap showed that it was no longer statistically significant when controlling for all the other measures in our model.

Predictors of attainment and their potential implications

While our findings support other established evidence showing that lower household income independently predicts lower attainment, they also suggest that other measures that may indicate material disadvantage, such as living in rented accommodation and/or a deprived neighbourhood, also play a part in explaining poorer results. These factors may not directly cause lower attainment – they may instead be linked to, or reflect, other cause and effect relationships. For example, the attainment penalty associated with living in rental accommodation may stem from a range of associated issues: the amount of space available for study in rented properties; the impact that insecurity of tenure may have on family dynamics; differences in the amount of neighbour noise, and so on.

The exact nature of these relationships cannot entirely be revealed by analysis of LSYPE2. While the data are rich and wide ranging, they have been collected

prospectively for broad analytical purposes – the surveys were not specifically designed to answer the research question 'What factors influence key stage 4 attainment?'. Given the need to ensure that LSYPE2 participants are not subjected to unacceptable respondent burden, this inevitably means that some parts of the picture are missing. For example, there is no data on the early development and cognitive ability of the LSYPE2 cohort. Furthermore, because of the wide scope of the analysis in this report, which examines many variables across the individual, school and area levels, in depth analysis of each of the specific variables which emerged as being important has been unfeasible at this stage.

Nevertheless, our analysis does show strong statistical relationships that are likely to signpost ways in which educational attainment could be improved. In addition to the material deprivation factors mentioned earlier, the findings confirm how circumstances relating to many aspects of a young person's life (including not living with parents, poor health, learning difficulties and poor psychological wellbeing) all tend to reduce attainment. These are well-recognised types of disadvantage where resources and services are already aimed at providing support. The schools system is designed to address some of these issues via specific policies such as the national funding formula, free school meals, the pupil premium, and reformed special educational needs and disability provision. There are also practices (formal and informal) within schools to support those who face the challenges of various types of disadvantage. But the fact that these issues continue to predict attainment deficits underlines the importance of ongoing efforts to address them.

Another important point to emphasise is that many factors which are predictive of attainment (low income, health, separation from parents etc.) span the remit of government departments and public services. As such, the educational detriment they are associated with cannot be completely addressed by better education policy or improved schools alone. More wide-ranging support for disadvantaged families – for example relating to income related benefits, childcare, children's centres, children's social care, and child and adolescent mental health services – is required to help ensure children have the best chance of fulfilling their potential at school.

Thanks to the breadth of LSYPE2 data, many of our findings relate to less well-recognised or less well-quantified types of educational disadvantage. New policies or revised practices could be effectively deployed here once other evidence or new research has helped to us to understand the causal mechanisms at play.

A number of findings suggest changes in home environment could improve education outcomes for some young people and merit further investigation. The answers to some of the following questions may already exist in published literature (given the scope of this report we have not conducted a full evidence review at this stage) while others may need further primary research to answer:

- Why are parents more inclined to encourage girls to pursue academic pathways than boys? Would making them aware of this bias and its effects improve boys' attainment? Are parental expectations affected by different levels of educational engagement between boys and girls?
- Given the link between attainment and maternal qualifications, would improving mothers' qualification levels during their child's lifetime improve the young person's own attainment? If so, at what point in their development would children benefit most from their mother re-engaging with education themselves?
- Why does a lack of a home computer connected to the internet lead to lower attainment? Where is this most problematic and could this issue be addressed by government and/or charities?
- Do parents and young people understand that a poor relationship between them affects attainment? Is there a role for school councillors or some form of mediation to help tackle problems at home that are less serious than child protection issues?
- Do parents and young people understand the importance of completing set homework or do they see it as a needless exercise? How can these perceptions be most effectively challenged?

A number of findings relate to young peoples' behaviour:

- Why is it that young people engaging in positive or prosocial behaviours (reading for pleasure, playing an instrument or sport, and/or community activity) have better attainment? Do these activities involve more education relevant learning than alternative leisure activities? Could participation campaigns help improve grades?
- Could truancy be tackled more effectively if young people, parents and teachers all realise the size of the educational penalty associated with it?
- Why isn't misbehaviour in class statistically linked to lower attainment when it is linked to less progress? Could understanding these mechanisms help identify how to encourage those children not reaching their full potential?

And some findings could have implications for schools, young people and their parents:

- Is it understood that the minority of young people who consider themselves to be 'picked on' by teachers, and who perceive that teachers try less hard with them, have lower attainment and progress? Young people also have lower progress if they don't like their teachers. Why do these perceptions arise and would a focus on teacher–pupil relationships improve attainment?
- Why is lower attainment predicted by both high *and* low levels of psychological distress? Is this because a degree of education-related anxiety can be beneficial to attainment? What are the longer term implications of any level of psychological distress, and how can we identify and guide young people who have too much?
- Why does the ethnicity of Black African and Black Caribbean pupils continue to predict a degree of lower attainment after other factors have been accounted for? What explanatory factors might be missing from our statistical models? Could primary and/or secondary schools take further action to support them?

The influence of the some of the factors identified above might be relatively small while the influence of others might be substantial but only relevant to a minority of young people. Such assessments of scale could be used to inform decisions about which of the issues outlined above (along with other questions which the findings in this report may suggest) are most pressing when it comes to any development of policy or practice.

The layering effect of disadvantage on attainment

Our data also underline the extent to which coming from a less privileged personal or family background is associated with attainment and, by extension, life chances. It also shows how disadvantage can manifest through many different circumstances, and how more nuanced analysis is required to understand which are most important and how they interact. Even after controlling for all other factors in our data, those living with neither of their biological parents had lower attainment than those living with both biological parents. Those whose mother held a degree level qualification had higher attainment than those with less qualified mothers. Those in below median income households performed less well than those with higher income. Those living in the most disadvantaged neighbourhoods had lower attainment than those in the most advantaged neighbourhoods. Various other such relationships are also illustrated throughout the report.

As shown in the illustration in Figure 41, if we assume, as supposed by our model, that the effects are linearly additive, a young person experiencing the seven specified elements of disadvantage will, on average, have attainment around 14 grades lower than our comparator group[100] for these measures (all other things being equal). In such a case, the layering effects of disadvantage would be pronounced – a difference of 14 grades equates to the difference between a student attaining 7 B-grade GCSEs and a student attaining 7 D-grades. And a young person experiencing further elements of disadvantage (for example living without either of their biological parents) may have even more pronounced attainment deficits. <u>It is important to note that this is an approximate estimate</u> because we do not test whether these effects are truly additive. For example, it is quite possible that certain combinations of disadvantage will have a greater or lesser effect than the sum of their constituent effects as reported here – this is a topic that we will investigate further in a forthcoming report on the attainment effects of multiple disadvantage. And, as with all data quoted in this report, there are margins of error around each of the figures in Figure 41, which may also cumulate[101]. It is also important

---

[100]The comparator group here is young people living in an owner-occupied household in the least deprived IDACI quintile with above median household income, a mother with degree level qualifications and an internet connected PC, who has not reported an illness or disability or being cyber-bullied in the last 12 months. Both disadvantaged and comparator groups are average across all other measures in the full model.

[101] This means, assuming the effects are linearly additive, the impact of having all seven disadvantages would be somewhere between 6 and 23 attainment grades overall (although it is also more likely to be closer to our estimate of 14 grades than at these margins)

to note that these are the effects of disadvantage after adjustment for many other factors, including many of the mechanisms through which disadvantage is assumed to impact on attainment such as through lower aspirations.

Figure 41 An illustration of the way in which the effects of disadvantage may cumulate

	Lives in private rented housing	and	Is eligible for Free School Meals	and	Lives in most deprived IDACI quintile	and	Doesn't have computer with internet access	and
If young person…								
Individual effect on attainment…	-16.1 points (~2.5 grades)		-12.2 points (~2 grades)		-10.0 points (~1.5 grades)		-11.4 points (~2 grades)	
Potential cumulative effect on attainment…	-16.1 points (~2.5 grades)		-28.3 points (~4.5 grades)		-38.3 points (~6.5 grades)		-49.7 points (~8.5 grades)	

Mother has no qualifications	and	Has been cyber bullied	and	Has illness or disability affecting school
-15.3 points (~2.5 grades)		-7.7 points (~1 grade)		-14.2 points (~2 grades)
-65.0 points (~11 grades)		-72.7 points (~12 grades)		-86.9 points (~14.5 grades)

These findings show how critical it is to continue efforts to ensure that being born and raised in difficult circumstances should not restrict young people to a life trajectory that results in unfulfilled potential and yet further deprivation during their adult and (for many) parental stages of life. The effect is clearly of relevance when it comes to addressing enduring social division, one of the 'five giant challenges' addressed by the government's manifesto[102].

Finally, it should also be noted that while improving educational attainment may be *necessary* to overcome disadvantage, it may not be *sufficient* to address the issue in itself. Broader policy actions may also be needed to ensure that disadvantaged young people are, for example, instilled with the confidence, motivation and broader life skills which may positively contribute to their education and progress through life. Ensuring that they also have access to the same opportunities as those from less disadvantaged backgrounds will also be of vital importance.

---

[102] The government's manifesto can be found here:
https://www.conservatives.com/manifesto

Appendix A Methodology

Introduction

The approach to the analysis used in this research is introduced at various points in the report. Rather than repeat this information here, we provide pointers to these sections and provide additional information in response to a series of questions we believe analysts will want to have answered.

There are also several areas that we think readers and analysts will want to explore themselves, so we have provided an Excel document that can be found at www.gov.uk/government/publications/understanding-ks4-attainment-and-progress-evidence-from-lsype2 which contains a series of spreadsheets. These spreadsheets are grouped into Appendices, as outlined below.

- Appendix B provides the bivariate associations between each of the variables used in this report and the two outcome measures, Best 8 and the Level 2 English and maths threshold. Since it was not possible to describe every bivariate relationship in Chapter 2, this allows interested readers to look at other measures of interest., linear (best 8) and logistic (Level 2 basics) regression, were used to assess statistical significance for each independent variable. Continuous variables such as level of unauthorised absences in school were recoded into quintiles.
- Appendix C provides the univariate descriptive statistics of all our measures for the full sample, the complete cases samples used in the modelling, and the excluded sample. Differences in the complete cases and excluded sample were assessed using (independent) t-tests. It also provides the modelling coefficients for the fullest available sample at each stage of the modelling process (the point at which each domain or set of measures was entered into the model) alongside the coefficients for the complete cases at the same stage. Information about the characteristics of the additional 101 individuals who are dropped from the analysis in Chapter 4 can be provided on request. The tables follow the approach used successfully elsewhere[103].
- Appendix D provides the results of the statistical models that form the basis of the findings presented in Chapters 3 and 4. An explanation of the different models is given within the spreadsheet Appendix Index.

---

[103] For example, see the supplementary material in Sacker A, Ross A, MacLeod CA, et al Health and social exclusion in older age: evidence from Understanding Society, the UK household longitudinal study J Epidemiology Community Health 2017;71:681-690

In Chapter 1 there is an overview of the LSYPE study, its methodology and the analysis approach taken in this study. We expand on this methodological overview in the sections which follow.

LSYPE as a data source

This research report is based on the responses of the second cohort of young people to be involved in the Longitudinal Study of Young People in England (LSYPE2). LSYPE2 is a robust and wide-reaching study which is managed by the Department for Education.

The first round of interviews for LSYPE2 took place in 2013 and participants have been re-contacted on an annual basis, allowing us to understand how their lives change over time. LSYPE2 builds on the first LSYPE cohort study, which initially ran from 2004 to 2010 with a follow-up conducted by UCL IOE in 2015/16 when the participants were 25 years old[104]. As with the first LSYPE, the Department for Education is supporting the second cohort of surveys of young people from the age of 13/14 to 19/20.

The aims of the study are:

- to follow a sample of young people through the final years of compulsory education;
- to follow their transition from compulsory education to other forms of education, training, employment and other activities;
- to collect information about their career paths and about the factors affecting them; and
- to provide a strategic evidence base about the lives and experiences of young people.

LSYPE2 has both strengths (described below) and weaknesses (discussed in the next section). Its strengths include:

- **Scale**: the data comes from a very large initial study which is randomly sampled from across England. It has high response rates and boost samples of key groups of interest, which makes it suitable for this kind of analysis and for looking at sub-groups with a low incidence amongst the wider population.

---

[104] Known as Next Steps. The Institute of Education are now responsible for this survey see: http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=1246&sitesectiontitle=Welcome+to+Next+Steps+(LSYPE)

- **Breadth**: the breadth of the topics asked about in LSYPE2 is wide-ranging, meaning that the modelling conducted throughout the report is able to take account of a large number of the factors which may impact on attainment. In this report we have focused on the following domains, each of which spans a number of more specific measures:

  - individual characteristics of the young person
  - their family background, including social and economic position
  - the home environment, including parental attitudes and support
  - the young person's attitudes and behaviours
  - the young person's health and wellbeing, including experiences of bullying
  - perceptions of peers and teacher relationships within school
  - measures of the structure, composition and quality of the school
  - characteristics of the area in which they live.

- **Multiple data sources**: data is not only collected from the young person but also from their parents, which allows us to link information about parental background, socio-economic and home environment, including parental aspirations and expectations about the young person's future. This survey data is also linked to aggregate information about the school that the young person attends, for example its size, level of unauthorised absences and its most recent Ofsted report level. In addition, it is linked to external data about the characteristics of the area in which the young person lives and attends school, drawn from Census data. Finally, consent was collected from young people and their parents in order to allow the Department for Education to link data recorded about them on the National Pupil Database (NPD) with their survey data, greatly enhancing the richness of the study. This is of particular importance in this report as GCSE and equivalent attainment data has been linked to the survey data for individuals who gave their consent. This means that we have objective information about the educational outcomes at KS4 for the majority of study participants. We have also linked KS2 data with a view to better understanding the progress made throughout secondary school. The predictive value of linked variables was tested using our standard approach i.e. data completeness; descriptive metrics against target variables (e.g. strength of correlation); and, most importantly, theoretical relevance.

- **Selection of individuals within schools within areas**: The LSYPE2 sample was selected to reflect the natural hierarchical structure in which students (level 1) are clustered within schools (level 2), which in turn are clustered within areas (level 3). Sampling in this way makes it possible to acknowledge that students within a given school are typically more alike and have more similar experiences than students selected from different schools, because school characteristics such as the

leadership style or level of disadvantage experienced by the student intake may influence the attainment of its students. Likewise, we can allow for the possibility that students within schools within a specific area are more alike than students within schools selected from different areas, because phenomena such as the overall level of disadvantage in the area may influence the attainment of the children within it. We take account of this using a statistical approach called multilevel modelling and discuss this further below. This helps us identify how far differences in pupil attainment are to do with pupils themselves, the schools that they attend, or the areas in which their schools are situated.

Limitations

LSYPE2 represents a rich source of data covering wide-ranging aspects of young people's lives. It should be noted however that some potential influencers of attainment were not included in our analyses. Examples of these potentially 'missing' data include: early childhood measures (given that LSYPE2 begins at Year 9); metrics relating to the young person's siblings; direct teacher feedback on the young person; measures of the ethos of the young person's school; and access to quiet space at home (e.g. for homework). This is an illustrative list and should not be viewed as comprehensive given the complexity of the factors which may impact on attainment. Nevertheless, LSYPE2, and the supplementary sources which have been linked to the survey data, constitutes a robust and wide-ranging source which has allowed us to conduct analyses that provide new insights into the sometimes hidden factors that correspond with attainment and progress.

As with all surveys, the figures presented in this report are subject to sampling error. Sampling error is the known level of imprecision in the results arising from the fact that the data are based on a sample of young people in England (albeit a large sample) rather than a census of *all* young people in England. The uncertainty around each estimate can be expressed in the form of error bars which are included in all charts in Chapter 2. In addition, confidence intervals (or margins of error) are presented for each estimate in the statistical models are presented along with sample sizes in Appendix D. The error bars around each figure show the values between which we are 95% certain the true population value will fall. The potential error in the estimations described by these confidence intervals follows a 'normal distribution', which simply means the true population score is far more likely to be close to estimate than at the edge of the confidence interval. This is the reason why estimates with overlapping confidence intervals may nevertheless be found to have a statistically significant difference from one another.

Throughout the report, wherever we highlight differences between specific groups, these are statistically significant unless otherwise stated, which means we are confident that a difference also exists in the wider population from which our sample is drawn[105]. However, whilst statistical significance is critical and tells us that observed differences are unlikely to correspond to randomly occurring patterns in the survey data, it is also important to consider whether this difference is substantively interesting. This is particularly the case where sample sizes are large, as this means that small differences or effects, which would not realistically have any bearing on policy, can still meet the threshold of statistical significance. As a result, we encourage readers to focus on the size of observed effects rather than statistical significance per se.

Our final caveat is that when discussing the statistical models we often talk about differences in attainment (such as girls doing better than boys) being "explained" by other measures (such as behavioural or attitudinal measures) which differ between the genders. We use the term "explain" in a statistical sense. For example, in the case above, we would mean the extent to which a given behavioural or attitudinal measure may account for some, or all, of the association between gender and attainment. Whilst it is true that this factor might also explain why girls have higher levels of attainment than boys in a real sense of the word, we often cannot infer this level of causality. To do that, we would need clear evidence about how the causal relationship actually works, which is not something we can easily do with an observational study of this kind or with this statistical approach. We might say that a research finding supports a particular hypothesis but this is not its absolute proof.

LSYPE2 research method

Interviews were conducted in respondents' homes. At Wave 1, in 2013, interviews were carried out with the young person and up to two parents (where two parents lived in the same household as the young person). At Waves 2 and 3 (in 2014 and 2015 respectively), the young person and a nominated main parent[106] were interviewed. Interviews for all waves took place between April and September of the given survey year. Consent to allow LSYPE2 data to be linked to NPD data was collected as part of the LSYPE2 interview.

At each wave, the same young people that responded at prior waves were interviewed. A response rate of 71.6% was achieved at Wave 1, 85.5% at Wave 2 and 89.0% at Wave 3.

---

[105] All such data are statistically significant at p<.05
[106] The main parent was nominated by a parent or guardian in the household before the interview, and was the parent/guardian most involved with the young person's education

Sample definitions

The LSYPE2 cohort sample

It is intended that LSYPE2 will track a sample of 13,100 young people in England from the age of 13/14 annually for seven years, through to the age of 19/20.

The target survey population for LSYPE2 (defined at Wave 1 and followed up in consequent waves) was pupils normally residing in England, who attended year 9 between September 2012 and August 2013 and turned 14 years old within that time-period. Pupils from the maintained sector were sampled directly from the National Pupil Database (August 2012 edition). To draw their sample, the National Pupil Database was stratified by: region; school type; the proportion of year 9 pupils eligible for FSM; the proportion of year 9 pupils with special educational needs; and the proportion of year 9 pupils who have white British ethnicity. Pupils in the independent sector or pupils exclusively registered at Pupil Referral Units (both not covered in the National Pupil Database) were sampled using a two-stage process. First, independent schools and Pupil Referral Units were sampled directly from Edubase, after stratifying by region. The smallest schools (with fewer than five year 9 pupils) were excluded from the sampling process for practical purposes. Pupils were then sampled directly from cooperating independent schools and Pupil Referral Units. We discuss the specific types of school included in our statistical modelling in the following section.

The reported sample

The survey data for this report are primarily drawn from the second wave of LSYPE2, when the young people in our cohort were in year 10, during the 2013/14 academic year. The age of respondents at the time of interview would have been either 14 or 15. GCSEs were generally linear in this cohort, without modules being assessed in Year 10. Year 10 behavioural and attitudinal questions are therefore the closest we have to understanding the young people's context before their final assessments while also avoiding the issues related to using concurrent predictors (which would apply to Year 11 data). Where necessary, data from earlier waves of LSYPE2 are included in our analyses. To help contextualise the findings in this report, Table 1Table 2 below summarises the phasing of the LSYPE2 waves.

Table 2 Age and timing of the LSYPE cohorts

Wave	School year	Academic year	Age (years)	LSYPE2 interview (year)
Wave 1	Year 9	2012/13	13/14	2013
Wave 2	Year 10	2013/14	14/15	2014
Wave 3	Year 11	2014/15	15/16	2015
Wave 4	Post-compulsory (year 12)	2015/16	16/17	2016
Wave 5	Post-compulsory (year 13)	2016/17	17/18	2017
Wave 6	Post-compulsory (potentially 1st year HE or gap year)	2017/18	18/19	2018
Wave 7	Post-compulsory (potentially 2nd year of HE)	2018/19	19/20	2019

The LSYPE2 data are supplemented with administrative data from the National Pupil Database which captures KS2 attainment data (from towards the end of the cohort's primary education) and KS4 (GCSE or equivalent) examination performances at the end of year 11. We also source a range of school-level data from the NPD, including variables such as the incidence of young people having a Special Educational Need (SEN) within schools, and draw upon other sources of administrative data where applicable (e.g. the Income Deprivation Affecting Children Index (IDACI) measure of area deprivation).

Although much of the data is taken from responses at Wave 2, this report is based on the sample of young people who participated in Wave 3 *and* consented to data linkage at that time. Because NPD data is not available for young people in the independent school sector, this study is based only on young people in schools in the maintained sector. As a result, our main sample which forms the basis of Chapter 2 in this report is of 9,076 individuals (compared to the total achieved sample size of 10,010 at Wave 3).

Table 3 School types included in analyses

Type of school attended	Sample
Individuals in-scope for all chapters	
Academy Converter	3,646
Community School	2,030
Academy Sponsor Led	1,411
Foundation School	996
Voluntary Aided School	629
Voluntary Controlled	129
Free Schools	22
Studio Schools	13
University Technical	11
City Technology College	3
Sub-total	8,890
Additional individuals included in Chapter 2 only	
Community Special School	101
Pupil Referral Unit	20
Academy Special Converter	15
Foundation Special School	8
Non-Maintained Special School	1
No school identifier available	41
Sub-total	186
Total individuals	9,076

Table 4 Base sizes for full sample for Chapter 2

	Variable	Category	N
Personal characteristics	Gender	Girls	4589
		Boys	4485
	Term of birth	Autumn	2996
		Spring	2953
		Summer	3116
	Ethnicity	White	6991
		Mixed	380
		Indian	203
		Pakistani	334
		Bangladeshi	217
		Black African	396
		Black Caribbean	274
		Other	274
	Primary language	English only	7771
		English first	768
		Bilingual	144
		Other first	392
	Illness/disability status	No illness or disability	7572
		Yes - no school effect	725
		Yes - affects school	759
	SEN status (including statements and/or school action)	Without SEN	7056
		SEN	1979
Family background	Family Type	Two biological parents	5148
		One biological parent and one step-parent	892
		One biological parent	2845
		No biological parents	187
	Proportion of parents in household who work	100% of parents in household in work	5747
		50% of parents in household in work	1556
		0% of parents in household in work	1667
	NSSEC	Managerial + professional occupations	3758
		Intermediate occupations	2117
		Routine and manual occupations	2823
		Never worked/long-term unemployed	301
	Tenure	Own/Mortgage/Part-own part-rent	5326
		Rented from Council/Housing Association	2607
		Private Rent	1032
		Other	86

	Variable	Category	N
	Eligibility for FSM and household income	Household income greater than median income	3221
		...less than median income & not currently eligible for FSM	2043
		Have been eligible for FSM in the last 6 years	1506
		Currently eligible for FSM	1950
	Highest maternal qualification	Degree or equivalent	1649
		Higher Education below degree level	1118
		GCE A Level or equivalent	1231
		5 or more GCSE at A*-C or equivalent	1633
		Level 1 qualifications/some GCSE passes or equiv.	1849
		Other qualifications	137
		No qualifications	1123
		No mother present in household	234
School	School size (No of pupils)*	Quintile 1 (24-711 pupils)	1793
		Quintile 2 (772-983 pupils)	1795
		Quintile 3 (984-1194 pupils)	1820
		Quintile 4 (1195-1448 pupils)	1815
		Quintile 5 (1449-2640 pupils)	1853
	Whether school has a sixth form	Sixth form	5942
		No sixth form	3089
	Whether school has a selective admissions policy	Selective	320
		Not selective	8570
	Whether school is single gender or co-educational	Single gender	1013
		Co-educational	8022
	School ethnicity	Quintile 1 (0-52.7)	1801
		Quintile 2 (52.8-79.8)	1778
		Quintile 3 (79.9-89.8)	1824
		Quintile 4 (89.9-94.0)	1774
		Quintile 5 (94.1-100)	1858
	School FSM*	Quintile 1 (0-7.0%)	1807
		Quintile 2 (7.1-11.0%)	1743
		Quintile 3 (11.0-16.7%)	1871
		Quintile 4 (16.8-26.0%)	1806
		Quintile 5 (26.2-66.8%)	1808
	School ESL	Quintile 1 (0-1.5%)	1742
		Quintile 2 (1.6-3.6%)	1869
		Quintile 3 (3.7-8.6%)	1804
		Quintile 4 (8.7-23.4%)	1813

	Variable	Category	N
		Quintile 5 (23.48-93.5%)	1807
	School SEN*	Quintile 1 (0-0.8%)	1652
		Quintile 2 (0.9-1.3%)	1751
		Quintile 3 (1.4-1.9%)	1970
		Quintile 4 (2.0-2.9%)	1819
		Quintile 5 (3.0-100%)	1813
	Unauthorised absences*	Quintile 1 (0-0.6)	1898
		Quintile 2 (0.7-0.9)	1749
		Quintile 3 (1.0-1.3)	1736
		Quintile 4 (1.4-2.0)	1871
		Quintile 5 (2.1-23.4)	1756
	Ofsted school rating	Outstanding	1954
		Good	4194
		Requires improvement	2239
		Inadequate	615
Area	Region	London	1298
		East of England	949
		South West	799
		South East	1401
		North East	468
		East Midlands	832
		North West	1275
		West Midlands	1050
		Yorkshire and Humberside	1004
	Income deprivation affecting children index (IDACI) quintile	Quintile 1 (advantaged)	1492
		Quintile 2	1584
		Quintile 3	1579
		Quintile 4	1915
		Quintile 5 (disadvantaged)	2506
	Urban/rural classification	Urban	7275
		Rural	1757

* The ranges shown for these variables are wider than the ranges presented in Appendix C because the sample here includes individuals with no school identifier or attending one of a number of types of special school

In the statistical models employed in Chapters 3 and 4 we exclude the 41 young people for whom the school attended was unknown and the 145 who attended special schools of different types since we did not have sufficient data or sample size to report on these students in this way. This reduced the in-scope sample size for these chapters to 8,890, as outlined in Table 3 above. The sample size for these chapters was reduced further because we included a large number of variables in the statistical models, some of which

had missing values. Our analysis consequently focused on 'complete cases', i.e. records of pupils for whom we had valid answers to all the measures included in the models[107].

This means that the analysis in Chapter 3 is based on 4,200 individuals. In Chapter 4, we introduce KS2 attainment data from administrative sources into the model which adds further complexity because this data is missing for a substantial proportion of the LSYPE sample (approximately 30%). The reasons for this are discussed in the introduction to Chapter 4. This issue has been addressed by using imputed KS2 data generated by a DfE funded project led by RAND Europe. Imputed data for KS2 were not available for every LSYPE individual so the models in Chapter 4 are based on 4,091 individuals.

How were the data weighted?

The process for computing the LSYPE Cohort 2 Wave 3 weights is detailed in the Wave 3 technical report. To account for the effect of excluding non-consenting Wave 3 respondents, the LSYPE Cohort 2 Wave 3 weights were adjusted by a weighting factor, which accounted for Wave 3 respondents' differential probabilities of consenting to data linkage. This weighting factor is equal to the inverse probability of linkage consent, estimated using a logistic regression model. These weights are used for the descriptive analysis in Chapter 2 and elsewhere in the report.

The multilevel modelling carried out in Chapters 3 and 4 demands a slightly different approach to account for these weights since the procedures used require an individual weighting variable for each level. The school and area level weights used were proxy weights. Effectively, school (level 2) and area (level 3) level weights are weighted averages of the individual-level weights at the appropriate (school or area) level.

Approach to modelling the data

Statistical analysis and modelling included in this report was carried out in the STATA 14.0 software package, with some exploratory analyses conducted using Mplus 7.3.

Chapter 2 is based on bivariate analysis of the sample of 9,076 young people in maintained schools described earlier. As noted above, Independent schools were excluded from our analyses because NPD data are not available for this group. As is typical of such analyses, the focus on mean differences and correlations can mask range

---

[107] An analytical alternative to the complete-cases analysis performed for this report would use the technique of multiple imputation to fill in gaps from missing data. However, this alternative was beyond the resources available for this project.

(both in the aggregated measures of attainment and the range within the young people's responses).

Chapters 3 and 4 are based on the reduced sample (excluding young people in special schools and those with missing data as described earlier) and are based on a regression technique called multilevel modelling which takes account of the hierarchical structure of the LSYPE study, where pupils (level 1) are nested within schools (level 2) within areas (level 3). By using this approach, we are better placed to understand the extent to which differences in attainment are driven by individual factors or by systematic differences in the profile of the schools they attend or the areas where they live.

We developed our models in stages, including separate domains of measures sequentially. The multilevel modelling approach makes it possible to examine how much variation is attributable to the individual, school and area level, and how much individual and school variation is explained by different domains. We acknowledge throughout that the effect of each domain will partly reflect the order in which the different domains are introduced; our choice of order was based on the literature reviewed (as described in the next section).

As discussed in Chapter 3, we trialled two-level and three-level models. It was decided that two-level models were the better option because, although significant, the amount of variation explained at the third level relating to area was fractional whilst the computational time taken to run all models with three levels was excessively large. The data in this report are therefore based on two-level models. We carried out sensitivity analysis on the final model and found that there was limited difference in the coefficients for the two and three level models (data available on request). Further information about multi-level modelling approaches can be found at https://www.cmm.bris.ac.uk/lemma/.

The approach taken in this report makes it possible to explore attainment and progress, but other approaches are of course possible and we encourage others to investigate and build on the findings presented here.

Variable selection

We included a large number of variables in the models, across eight main domains (personal characteristics, family background, home environment, attitudes and behaviours, health and wellbeing, school and area factors). We based the selection process on an initial review of the existing literature in this field identifying a theoretical basis for the different domains and for the ordering of domains (beginning at the individual level, then school and area, and within individual beginning with more structural

factors)[108]. We cross-referenced these domains to the variables that are collected in LSYPE and then examined each variable for missing data, its bivariate relation to Best 8, and constructed summative scales where there was evidence of an underlying latent constructs (e.g. school attitudes and locus of control).

What is the effect of using a completed cases approach?

Our analysis uses a complete cases approach to dealing with missing data, which is the most commonly used approach in analyses of this type. Here we examine the potential impact on our findings.

A comparison of the fullest sample[109] with the complete cases sample across each measure suggests that differences in the distribution of our sample resulting from missing data are relatively small. However, where differences do exist they tend to show our analysis sample as less disadvantaged than in the general population. For example, whereas the proportion of young people living in routine and manual households is 22.0% in our analysis sample (Appendix C, column G), the respective figure in the fullest sample is 27.7% (column E). This means that factors associated with disadvantage are also a little under represented. For example, those not likely to, or not knowing whether they would apply to university is 8.0% in our analysis sample compared to 12.2% in the fullest sample. This is also evident in a statistical comparison of the differences between those included and excluded in our final analysis sample shown in column K.

Overall, differences are relatively small. Nevertheless, an issue arises where the sample size for categories of individuals are especially small and identification of a statistically significant difference is therefore more difficult. Typically, these categories represent especially disadvantaged subgroups of the population. Overall therefore, we are likely to underestimate the importance of disadvantage – our estimates of the strength of the relationships between elements of disadvantage and attainment and/or progress will tend to be on the conservative side.

As a further assessment of the potential impact of missing data we also examined differences in the model coefficients for each domain. This involved comparing the fullest sample (i.e. the largest available sample at the stage in which the domain was first entered into the model) with the complete cases sample at this same stage. Whilst some differences in the coefficients are evident, as would be expected, generally where it

---

[108] See References, particularly Chowdry, H., Crawford, C.& Goodman A (2009, 2010), Duckworth, K., R. Akerman, L. Morrison Gutman & J. Vorhaus (2009), Goodman, A. and Gregg G. (2010), Henderson, M. (2012), Khattab, N. (2014), Sutherland, A., Ilie, S. and Vignoles, A. (2015) and Webber, R and Butler, T (2007)
[109] The term 'fullest' is used instead of full sample because these sample sizes vary across the measures due to variable specific missingness.

matters (in instances where associations are statistically significant), differences are small, and the conclusions remain the same. However, there is some evidence of the underestimation of disadvantage. For example, the coefficient for living in a routine or manual household is b. -18.5 in the fullest sample (Appendix C, column L) and b. -11.8 in our analysis sample (column R). There is also evidence of an impact on effects associated with very small subgroups. For example, in the fullest sample the effect of living in a household where the NS-SEC is never worked/long-term unemployed is b. -16.3 (p = 0.025), whereas there was no association in our analysis sample.

Our assessments of the impact of missing data allow us to be reasonably confident that its impact is relatively small and predictable in terms of the underestimation of the impact of disadvantage and the factors associated with it. One way to assess this further would be to reanalyse the data using multiply imputed datasets to account for missingness and compare findings; something that we or other researchers would be able to do given additional resource.

References

Buscha, F. and Conte, A. (2014) The Impact of Truancy on Educational Attainment During Compulsory schooling: a bivariate ordered probit estimater with mixed effects, The Manchester School Vol 82 No 1 (103-127) doi: 10.1111/manc.12002

Chowdry, H., Crawford, C.& Goodman A (2009), Drivers and barriers to educational success: evidence from the Longitudinal Study of Young People in England, DCSF Research Report No. RR102

Chowdry, H., Crawford C, & Goodman, A. (2010), The role of attitudes and behaviours in explaining socio-economic differences in attainment at age 16, IFS Working Paper 10/15.

Crawford, C., Deardon, L. and Greaves, E. (2013) When you are born matters: Evidence for England, IFS Report R80, London

Demark, S., Platts-Fowler, D., Robinson, D., Stevens, A., and Wilson, I. (2010) Young People and Community Cohesion: Analysis from the Longitudinal Study of Young People in England (LSYPE) DfE Research Report No. RR033

Department for Education (2015) key stage 2 to key stage 4 Value Added Measures: A technical guide for local authorities, maintained schools, academies and free schools, DFE-90015-2015 http://dera.ioe.ac.uk/25969/

Department for Education (2016a) Revised GCSE and equivalent results in England, 2014 to 2015 SRF 01/2016, 21 January 2016 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/494073/SFR01_2016.pdf

Department for Education (2016b) Outcomes for children looked after by local authorities in England, 31 March 2015, SFR 11/2016, 24 March 2016 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/509965/SFR11_2016_SFRText.pdf

Duckworth, K., R. Akerman, L. Morrison Gutman & J. Vorhaus (2009), "Influences and leverages on low levels of attainment: a review of literature and policy initiatives", Centre for Research on the Wider Benefits of Learning, Institute of Education, University of London

Goodman, A. and Gregg G. (eds) (2010) with Chowdry, H., Crawford C, Goodman, A., Deardon, L., Joyce, R., Sibieta, L., Sylva, K., Washbrook, E. (2010) Poorer children's educational attainment: how important are attitudes and behaviour? IFS

Green, R. (2010) Educational Attainment Among Young People in Rural England – A Multilevel Analysis, ICLS Occasional Paper 4.2

Green, R., Collingwood, A. and Ross, A. (2010) Characteristics of bullying victims in schools DfE Report – RR001 ISBN: 978-1-84775-753-1

Henderson, M. (2012) A test of parenting strategies, Sociology 47(3) 542-559

Khattab, N. (2014) How and when do educational aspirations expectations and achievement align? Sociological Research Online, 19(4), 7 DOI: 10.5153/sro.3508

Lessof. C, Ross, A., Brind, R., Bell E. and Newton, S. (2016) Longitudinal Study of Young People in England cohort 2: health and wellbeing at Wave 2, DfE Research Report 501 ISBN 978-1-78105-561-8
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/599871/LSYPE2_w2-research_report.pdf

Saunders, C. L., Brunton-Smith, I., Garrod, B., Stepanek, M., Vignoles, A. and Sutherland, A. (2016) Missing key stage 2 data in LSYPE2: technical report, DFE-RR608, ISBN: 978-1-78105-669-1

Strand, S. (2007) Minority ethnic pupils in the longitudinal study of young people in England (LSYPE); available from
http://www.dcsf.gov.uk/rsgateway/DB/RRP/u014850/index.shtml

Strand, S. (2008) Minority ethnic pupils in the longitudinal study of young people in England: Extensions report on performance in public examinations at age 16; available from http://www.actiononaccess.org/resources/files/resources__DCSF-RR029.pdf

Strand, S. & Winston, J. (2009). Educational aspirations in inner city schools. Educational Studies, 34 (4). 249-267Strand, S. (2014). School effects and ethnic, gender and socio-economic gaps in educational achievement at age 11. Oxford Review of Education, 40(2), 223-245. doi:10.1080/03054985.2014.891980

Sutherland, A., Ilie, S. and Vignoles, A. (2015) Factors associated with achievement: key stage4, DfE Research brief 407 November 2015

Webber, R and Butler, T (2007) Classifying pupils by where they live: how well does this predict variations in their GCSE results *Urban Studies*, Vol 44 No 7 1229-1253 June 2007

Wigfield, A. & Eccles, J. (2000) Expectancy-Value Theory of Achievement Motivation, Contemporary Educational Psychology, Vol. 25, pp. 68-81

Any enquiries regarding this publication should be sent to us at: team.longitudinal@education.gov.uk or www.education.gov.uk/contactus

This document is available for download at www.gov.uk/government/publications