Measuring the disadvantage gap in 16-19 education

Sam Tuckett, David Robinson and Felix Bunting

March 2021

Research Area:
Higher Education,
Further Education,
and Skills

About the authors

Sam Tuckett is a senior researcher at the Education Policy Institute. Prior to joining, Sam worked for five years as an analyst in the Department for Education where his roles included analysis of key stage 2 and 4 accountability reforms and supporting the introduction of Progress 8. More recently he has worked across government to link education data sets with sources held by the Department for Work and Pensions and Her Majesty's Revenue and Customs to expand the existing evidence base. Sam holds a first-class undergraduate degree in Mathematics from Cardiff University.

David Robinson joined the Education Policy Institute in June 2017. David's background includes six years at the Department for Education as lead analyst first on school and college accountability and then on capital funding. Since joining EPI David has authored the following reports: '16-19 education funding: trends and implications'; 'Post-18 education and funding: options for the government review'; 'UTCs: are they delivering for young people and the economy?'; 'Further Education Pathways: Securing a successful and healthy life after education'; 'An international comparison of technical education funding systems: What can England learn from successful countries?'; and 'The impact of interventions for widening access to higher education: a review of the evidence'.

Felix Bunting is a researcher currently focused on post-16 and skills at the Education Policy Institute. Felix first joined EPI through the Undergraduate Research Internship programme in summer 2018, before returning post-graduation, in 2019.

Over the course of his studies, he interned as a quantitative analyst within industry, market research and public sector consultancy.

Before joining EPI, Felix studied Physics at the University of Oxford. He graduated with an MPhys in 2019, with his Master's project focused on the theoretical basis of machine learning.

Acknowledgements

About the Education Policy Institute

The Education Policy Institute is an independent, impartial and evidence-based research institute that promotes high quality education outcomes, regardless of social background. We achieve this through data-led analysis, innovative research and high-profile events.

Education can have a transformative effect on the life chances of young people, enabling them to fulfil their potential, have successful careers, and grasp opportunities. As well as having a positive impact on the individual, good quality education and child wellbeing also promotes economic productivity and a cohesive society.

Through our research, we provide insight, commentary, and a constructive critique of education policy in England – shedding light on what is working and where further progress needs to be made. Our research and analysis span a young person's journey from the early years through to entry to the labour market.

Our core research areas include:

- Benchmarking English Education
- School Performance, Admissions, and Capacity
- Early Years Development
- Social Mobility and Vulnerable Learners
- Accountability, Assessment, and Inspection
- Curriculum and Qualifications
- Teacher Supply and Quality
- Education Funding
- Higher Education, Further Education, and Skills

Our experienced and dedicated team works closely with academics, think tanks, and other research foundations and charities to shape the policy agenda.

Contents

Foreword

One of the greatest challenges facing England's education system is the large attainment gap between disadvantaged pupils and the rest of the pupil population. For many years, the size and trends in this gap have been carefully measured, including in EPI's Annual Reports. This existing evidence shows that the gap grows larger in each phase of education up to the end of key stage 4.

However, much less has been known about the size of and trends in the gap in the 16-19 phase of education. This is because a sound methodology has yet to be agreed upon to allow consistent and reliable measurement of this gap. The challenge of developing a robust 16-19 gap measure comes in large part as a consequence of the challenge of measuring the gap across a much wider and less easily compared set of post 16 qualifications. The challenge also relates to some young people leaving education after age 16 or pursuing routes such as apprenticeships.

Thanks to support from the Nuffield Foundation, this report seeks to address this challenge, by providing a new measure of the 16-19 disadvantage attainment gap. We are very grateful not only to the Nuffield Foundation but to those on our advisory group who have contributed to this project. We believe that the methodology set out in this paper does represent a sound way of measuring the 16-19 disadvantage attainment gap and will help policy makers to track this key metric and take associated policy decisions.

But we are keen to encourage all those with an interest in this issue to comment on the proposed methodology, so we can consider if there are any further changes that should be made before we adopt this as our preferred metric for measuring the 16-19 disadvantage attainment gap.

The 16-19 phase of education has often appeared to be the "Cinderella" phase of education, both in terms of research and policy attention. We hope that this report will make some contribution to remedying this relative neglect.

Rt. Hon. David Laws, Executive Chairman, Education Policy Institute

Executive summary

Disadvantaged school-age children tend to have worse attainment than their peers. This is well understood, and schools are funded to provide extra support for these pupils. Much less is known about the attainment of disadvantaged students during the final phase of compulsory education and how this compares to their peers. Using students' free school meal status during their last six years of school as an indicator of disadvantage, and a 16-19 attainment measure based on the qualifications and grades they achieved between the age of 16 and 19, we have undertaken new exploratory research and present our provisional methodology to fill this critical evidence gap.

There is a gap equivalent to almost three A level grades between the best three qualifications of disadvantaged students and their peers

In terms of all the qualifications achieved between the end of secondary school and by the age of 19, the average disadvantaged student is the equivalent of around three A level grades behind their non-disadvantaged peers. There was little change in this gap between 2017 and 2019, prior to the Covid-19 pandemic. For students identified as being disadvantaged over a prolonged period, this gap increases to the equivalent of almost four A level grades.

Disadvantaged students take fewer qualifications and are more likely to take vocational qualifications

Disadvantaged students are more likely to take vocational and technical qualifications. They also tend to enter fewer, and lower level, qualifications. However, in recent years the total number of qualifications entered has dropped by a greater amount for non-disadvantaged students than it has for disadvantaged students.

Disadvantage attainment gaps are largest within academic qualifications, such as A levels

Within the most common types of qualifications entered, disadvantaged students on average achieve lower results, with the largest gaps being in academic qualifications such as A levels. Disadvantaged students on average achieved half a grade less in every A level entered compared to non-disadvantaged students. This factor, combined with the fact that disadvantaged students are also likely to enter fewer qualifications, is a key driver of the size of the gap.

For non-academic qualifications that are equivalent to A levels, such as level 3 BTECs with similar teaching hours, this disadvantage attainment gap is only the equivalent of one quarter of an A level grade per qualification.

Disadvantaged students have lower average attainment during 16-19 study than otherwise similar students

To provide a greater understanding of 16-19 attainment and the associated gap, we have produced a number of statistical models to estimate the relationships between this attainment and a range of different student and institution characteristics. Students' prior attainment, the ability of their peers and qualification type appear to be key predictors of overall attainment. Even after controlling for these, and other student and institution characteristics, disadvantaged students still have lower 16-19 attainment than other students. This remaining gap is equivalent to almost half an A level grade.

The use of simple aggregate statistics to compare the 16-19 disadvantage attainment gap to that at key stage 4 must be treated with caution, as conclusions depend upon the metric used. If the disadvantage attainment gap is calculated as the difference in mean attainment between disadvantaged and non-disadvantaged students, it appears slightly narrower in the 16-19 phase compared to key stage 4. However, if the gap is measured as the difference in median attainment between the two groups it appears wider in the 16-19 phase compared to key stage 4. These differences arise largely as a result of the high proportion of students with zero attainment in the 16-19 phase (those that did not enter, or did not achieve a pass grade in any level 1 to 3 qualifications). Unless stated otherwise, we have presented figures based on a mean average throughout this report.

Although comparisons between phases using simple aggregate statistics should be treated with caution, the finding from our statistical modelling does suggests that disadvantaged students fall further behind during the 16-19 phase, when compared with otherwise similar students. That is, that student disadvantage is negatively associated with 16-19 attainment, even when controlling for GCSE results.

The lower prior attainment of disadvantaged students is the main factor behind the disadvantage attainment gap

Eighty six per cent of the gap between the 16-19 attainment of disadvantaged and non-disadvantaged students can be explained by other differences in the characteristics of the two groups. Of these characteristics, prior attainment explains 39 per cent of the total gap, the average prior attainment of students' peers explains 12 per cent and the type of qualifications entered explains 33 per cent.

One seventh of the disadvantage attainment gap is unexplained by other student or institution characteristics

Fourteen per cent of the disadvantage attainment gap cannot be explained by student or institution characteristics, equivalent to almost half an A level grade. This could be the continued effect of disadvantage itself, and/or it could be due to differences in unobserved characteristics such as health or motivation. Our regression models show that 92 per cent of the variation in attainment is due to the differences between students, so it is likely that this unexplained gap is associated with differences between students rather than institutions.

Disadvantage attainment gaps are most prominent for white British students

Disadvantaged white British students have amongst the lowest attainment, with disadvantaged students of most other ethnicities attaining more highly. Some of the difference in attainment at 16-19 seen by ethnicity can be attributed to differences in prior attainment, as well as other demographic characteristics. Once these factors are considered, the gaps between different groups shrink significantly. Only disadvantaged Gypsy/Roma and Traveller of Irish Heritage students have lower attainment than disadvantaged white British students.

16-19 attainment varies by the type of institution attended

Amongst the main types of provider, students at Further Education colleges have notably lower attainment than average. This appears to be largely related to differences in known student

characteristics, mainly prior attainment and type of qualification entered. After controlling for these factors, the difference reduces to the equivalent of just one seventh of an A level grade compared to local authority-maintained schools.

Area level and student level disadvantage measures appear similar in their ability to predict attainment

The area-based index of multiple deprivation (IMD) is currently used as part of the 16-19 funding formula, but student level disadvantage is not. Student level disadvantage (free school meals) measures explain marginally more of the variation in attainment than use of the IMD alone, and there are further marginal increases from including both measures. This implies that there are some disadvantaged students not targeted by the existing funding formula, and yet this disadvantage is likely to be contributing to their lower attainment.

The disadvantage attainment gap varies considerably across local authorities

In Knowsley, North Somerset and Stockton-on-Tees the 16-19 disadvantage attainment gap is over 4.5 A level grades, whilst in Southwark, Redbridge, Ealing, Sutton, Merton, Newham and Islington disadvantaged students have higher average attainment than non-disadvantaged students nationally.

Conclusions and implications for policy

The research undertaken in this report is exploratory. To highlight the overall status of disadvantaged students in the 16-19 phase, we have had to make assumptions about the equivalencies of different qualifications. However, even under our more conservative assumptions our analysis has shown that economic disadvantage could be leading to disadvantaged students dropping almost half an A level grade when compared to their more advantaged peers. For progression to further or higher education, or transitions to the labour market, differences of this size matter.

However, the majority of the gap between disadvantaged students and their peers appears to be explained by other factors, with two fifths explained by students' prior attainment. Beyond prior attainment, other factors make a smaller contribution, but nevertheless warrant further investigation. Both the ability of students' peers and the type and level of qualifications entered make a noticeable contribution to 16-19 attainment and the corresponding gap. This finding is consistent with existing research which shows high achieving disadvantaged students are also more likely to take qualifications associated with lower prior attainment (academic mismatch). Information, advice and guidance targeted at high attaining disadvantaged students should play a role in addressing this mismatch.

So how should policymakers respond to a gap that is mostly, but not completely, driven by prior attainment?

Whilst much of the focus should be on earlier phases, for the disadvantage attainment gap to close, a concomitant increase in efforts to limit the impact of disadvantage during the 16-19 phase is required. If disadvantaged young people are to avoid falling yet further behind, addressing this gap should be central to the government's reform agenda for the 16-19 phase and for further education.

Our findings also strengthen the case for including student level disadvantage measures within the 16-19 funding formula, alongside the area-based measures currently used. Introducing such funding as a Student Premium, alongside the associated accountability and transparency requirements for providers, would help heighten the focus on disadvantaged students during this phase.

Critically, these results also predate the Covid-19 pandemic and the resulting lost learning and disruption to exams; factors which may have exacerbated the disadvantage attainment gap. To ensure that existing and emerging inequalities are identified and addressed we will continue to review and refine the provisional methodology presented in this report and monitor the 16-19 disadvantage attainment gap through 2020 and beyond.

1 - Introduction

There is a large body of research considering attainment gaps between the most economically disadvantaged children and all others up to the age of 16. This allows for evidence-based policy making in these phases of education, for example the distribution of additional, targeted funding to schools with children from disadvantaged backgrounds. Furthermore, time series measures allow for monitoring of the disadvantage attainment gap to help evaluate the impact of specific policies or economic events.

Socio-economic differences in attainment for post 16 education are much less well understood.

Existing research into disadvantage in the 16-19 phase of education is focused primarily on choice of qualification and institution type, or binary measures of whether a certain outcome was achieved.

For example, the Centre for Vocational Education Research (CVER) found that students that were eligible for free school meals were under-represented in those that entered academic qualifications in the 16-19 phase, or held any level 3 qualification by age 20 (Hupkau et al. 2017). Furthermore, that those studying towards level 2 qualifications at age 17 were not likely to enter higher level qualifications in later years.

Further research from CVER shows that the peer effect (ability of students in the same school cohort) is a determinant of whether students are likely to pursue academic or vocational qualifications at age 16, beyond the effect of their own prior attainment. In addition, the research shows that this effect is stronger when those peers also have similar socio-economic backgrounds to their own (Hedges, Speckesser, and Lazarowicz 2017).

Using LYSPE (Longitudinal Study of Young People) CVER research finds that the majority of students who aspired to pursue an academic or vocational route at the end of compulsory education (which at the time, was age 16) did so. However, of those that did not aspire to continue post 16 but in fact did, most opted for a vocational route (McIntosh 2019).

Analysis from the Fisher Family Trust (FFT) data lab, shows that BTECs are potentially overvalued in a point score system with no adjustments (Thomson 2018). Students from disadvantaged backgrounds are more likely to enter vocational qualifications (including BTECs) so the impact of still popular, non-reformed qualifications being over valued could be to make the disadvantage attainment gap appear narrower. However, we present an alternative methodology in this report that adjusts the weighting of academic and vocational qualifications based on subsequent earnings, which will reduce the impact of any potential overvaluing.

With the participation age now set at 18 and against the backdrop of a recession, the lack of research into the 16-19 attainment of disadvantaged students has become far more prominent. New statistics, capable of regular update will inform policy intervention in this area and are a pre-requisite to making progress in closing the disadvantage attainment gap in this phase.

One of the difficulties in measuring the disadvantage attainment gap for 16-19-year olds compared to school age pupils, is the wider range of qualifications learners may choose from. Indeed, it is now the case that the majority of 16-19 students take vocational, technical or lower level qualifications, rather than A or AS levels. Existing statistics allow us to look at the differences in the grades achieved between disadvantage and non-disadvantaged students within each qualification type. However,

given not all students are equally likely to enter certain types of qualification, this does little to inform us as to how big the gap is overall and whether any progress towards closing it is being made.

This report presents a provisional methodology for new measures to monitor the 16-19 disadvantage attainment gap through time. It considers the difference in outcomes within each of the main qualification pathways available to learners at this phase. In addition, we also present a provisional composite measure including all level 1 to 3 qualifications for the first time.

Beyond looking to measure the disadvantage attainment gap, we also consider the factors most strongly associated with attainment at 16-19. For example, institution level factors such as school or college type, as well as individual characteristics such as gender and ethnicity. This research examines how much of the variation in results can be explained by disadvantage status once these other factors are held constant. We also go a step further and model the attainment students from economically deprived backgrounds may have achieved, if all of their other characteristics were similar to other students.

2 - Methodology

This report presents an exploratory analysis of the 16-19 disadvantage attainment gap. The methodology presented in this section and findings which follow should be considered provisional. EPI will continue to develop this measure and we welcome feedback on this methodology.

Coverage and data

All students at the end of their 16-19 study at a state-maintained school or college (other than those on apprenticeship programmes) have been included in this analysis.

Most students will enter this phase of education at age 16 and complete a two-year course, in years 12 and 13. There are also a minority of students who enter a three-year programme of study and complete this phase at age 19, in year 14.

We consider regulated qualifications at level 1, 2 or 3 that these students completed during their 16-19 study.

Students that appeared in key stage 4 data but did not appear in data indicating they had completed 16-19 study by age 19 (i.e., those that did not continue in any form of education beyond the age of 16) are not included.

Unless otherwise stated, all analysis presented in this report is derived from the National Pupil Database (NPD) student and exam files and the Ofqual qualification register. The National Pupil Database files are made available to EPI by the Department for Education via the Office for National Statistics Secure Research Service. The NPD covering this phase of education is a combined dataset covering students in school sixth forms as recorded in the school census, and students at other further education institutions as recorded on the Individualised Learner Record (ILR). The datasets, including student level information and information on the qualifications they have taken, are compiled by the Department for Education. The only additional matching completed within EPI is to link back to school census records to obtain information on free school meal eligibility.

Defining disadvantage

Unless otherwise stated, disadvantaged students are defined throughout this report as those who were known to be eligible for and claiming free school meals in any of the six years prior to finishing key stage 4. There are a small number of students who will have been eligible for a free meal but opted not to claim it who will not be covered by this definition. Although further education institutions have had an obligation to provide a free meal to students since 2015, take up is low and the data are not made available on the National Pupil Database. Students' free school meal history from when they were school aged is therefore the preferred measure of disadvantage.

In the latter sections of this report which present our regression modelling, we also refer to persistent disadvantage, by which we mean those that were eligible for free school meals for over 80 per cent of their time in primary and secondary school. We also consider the association between the area-based index of multiple deprivation and 16-19 attainment, and explore further the effectiveness of our main disadvantage definition in the 'factors associated with 16-19 attainment' section of this report.

Table 1: Students at the end of 16-19 study included in the analysis in each year

Year	Disadvantaged	Non-disadvantaged	Total students
2017	119,980	385,178	505,158
2018	120,049	378,839	498,888
2019	119,497	378,044	497,541

Table 1 shows the number of students included in the overall national measure for each year. Unless otherwise stated, these are the student numbers underlying any national analysis.

Students which could not be linked to a school census record, which provides information on whether they were eligible for free school meals at the end of key stage 4, have been categorised as non-disadvantaged for the purpose of this analysis.

Measuring the 16-19 disadvantage attainment gap

For this report, the 16-19 disadvantage attainment gap has been calculated at a national level as the mean average, equivalent number of A level grades that disadvantaged students were behind non-disadvantaged students, over their best three qualifications taken at level 1 to 3 in this phase.

If there were no difference on average between the attainment of disadvantaged and non-disadvantaged students, the difference in mean attainment between the two groups would be zero.

Measuring attainment

Before the disadvantage attainment gap can be measured, we must first define an attainment measure for each student, in the form of a total point score.

How points are allocated to different qualifications and grades will depend on what values are ascribed to qualifications, and there is no methodology that serves all purposes. This is especially the case for the 16-19 phase, given the multitude of pathways students progress onto afterwards e.g. apprenticeships, higher education, employment, all of which will have different qualification requirements.

After consulting experts across the sector, we have opted to present two methods, each of which applies different assumptions. We have further completed sensitivity testing around these assumptions to ascertain their impact. Under both methods, we measure the total point score from a student's best three qualifications, as this is the number of qualifications most commonly used for admission to further education and of interest to employers.

Method 1 – qualifications of equivalent level and teaching hours get equal points

Under the first approach, our underlying principle is that hours of teaching time (or educational inputs) for qualifications of the same level should be given equal value. We therefore assume that the same weighting should be given to all qualifications at the same level if they have a similar number of guided teaching hours.

For example, level 3 qualifications with similar teaching hours to an A level should be awarded the same points as an A level. Further detail as to how this approach is implemented is provided in Annex A and C.

Method 2 – method 1 adjusted for the higher labour market returns to academic qualifications

Under our second approach the underlying principle is that qualifications that deliver similar returns in the labour market should be given equal value. We use existing research (Patrignani, Battiston, and Conlon 2019) to apply a simple adjustment to increase the points of level 3 academic qualifications proportionately to the economic returns seen for students entering A levels compared to BTECs. More complex adjustments, for example to account for the difference in returns between subjects, would also be possible and we will look to further develop this adjusted methodology in future.

More broadly, we assess how sensitive the outputs are to different assumptions of equivalence between academic and vocational qualifications in Annex A.

Level 1 & 2 qualifications

In addition to the relative weighting of academic v vocational/technical qualifications, we also apply a weighting for level 1 and 2 qualifications relative to level 3 qualifications. We use the fact that an AS level is often taken alongside level 2 qualifications at key stage 4, or alongside other level 3 qualifications post 16. AS levels therefore have established point systems relative to both a 9-1 GCSE and an A level. This means AS levels can be used to create a broad mapping between the two. This methodology is described in more detail in Annex C. How sensitive the results are to these assumptions is also tested in Annex A.

Treatment of English and maths resits

Since 2014, those who did not achieve a GCSE pass in English and maths by the end of key stage 4 are required to continue study of these subjects as part of their 16-19 programme. This means that many students have entered a GCSE (or other approved English or maths qualification) in the 16-19 phase. However, only students achieving a grade higher than their key stage 4 grade will see the benefits from (re)taking this qualification. As such, for students that were required to continue English or maths study, we have included points associated with the improvement of their grade rather than points associated with the grade itself. If they achieved the same or a poorer outcome in their 16-19 resit compared to in key stage 4, we have not counted the qualifications at all.

For example, students achieving a 3 in key stage 4 and a 4 during the 16-19 phase will achieve a 1, and this will be aggregated into their 16-19 attainment score. If the same students instead achieve a 3 again during their 16-19 study, we will not include their GCSE resit. We have not taken this approach for other qualifications taken in key stage 4 and resat during the 16-19 phase, as only GCSE English and maths resits are (effectively) mandated by the government. There are also a substantial number of other GCSE and similar level one or two qualifications taken during 16-19 study, so to not include these would skew the measure considerably.

Regression models

To provide a greater understanding of the association of different student and institution characteristics with 16-19 attainment, we have produced a number of ordinary least squares (OLS) regression models. We have fitted a number of models, all with 16-19 attainment as the dependent variable, and student and institution characteristics as the independent variables. We have built them up iteratively to allow us to examine how the coefficients for different variables change, as an increasing number of factors are controlled for.

We have built upon our OLS models with multi-level models that recognise that students are clustered together within institutions. These models allow us to consider what proportion of variance in 16-19 attainment we can attribute to between institution variance, and within institution variance.

The between-institution variance is the amount of variation in attainment accounted for by the differences between schools and colleges attended, once other factors have been taken into account. The within-institution variance is the remaining proportion of variation, attributed to student level characteristics.

Oaxaca Blinder decomposition

To further analyse the relationship between disadvantage and the 16-19 attainment gap, we implement a technique known as the Oaxaca Blinder decomposition. This technique is a decomposition of the difference in attainment between disadvantaged and non-disadvantaged students into a component that is explained by differences in characteristics between the two groups, and a component that is unexplained by differences in these characteristics. For example, how much of the disadvantage attainment gap can be explained by the lower prior attainment of disadvantaged students.

Additional analysis using longitudinal survey data

Though the NPD and ILR are useful in their broad coverage of the population of 16-19 year olds, it is limited in the depth of information on each student. To look at a broader set of factors we make use of Next Steps, Longitudinal survey data[1] which has followed the lives of around 16,000 people born in 1989/90. We use this data to provide richer information such as future aspirations and level of parental education.

The data from this survey have been matched to National Pupil Database data, so that for 4,297 individuals, we can link to a 16-19 point score outcome for the 2006/07 academic year.

The point score measure is different from the one we present in this report under method 1 and 2, as it uses all qualifications, rather than a total across students' best three qualifications. In addition, it does not include qualifications below level 3. The point scores are those that were used by the Department for Education at the time and include all eligible qualifications that a student entered between the age of 16 and 18. Although the total point scores are not directly comparable, the relative value of qualifications and intervals between grades is broadly consistent with that of our analysis of more recent data.

---

[1] **Source:** University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2020). Next Steps: Linked Education Administrative Datasets (National Pupil Database), England, 2005-2009: Secure Access. [data collection]. 6th Edition. UK Data Service. SN: 7104, http://doi.org/10.5255/UKDA_SN_7104-6

3 - The 16-19 disadvantage attainment gap

Figure 1 below shows the disadvantage attainment gap in A level grades between 2017 and 2019, shown for both method 1 and method 2 as described in the previous section.

Figure 1: 16-19 Disadvantage attainment gap - method 1 and 2
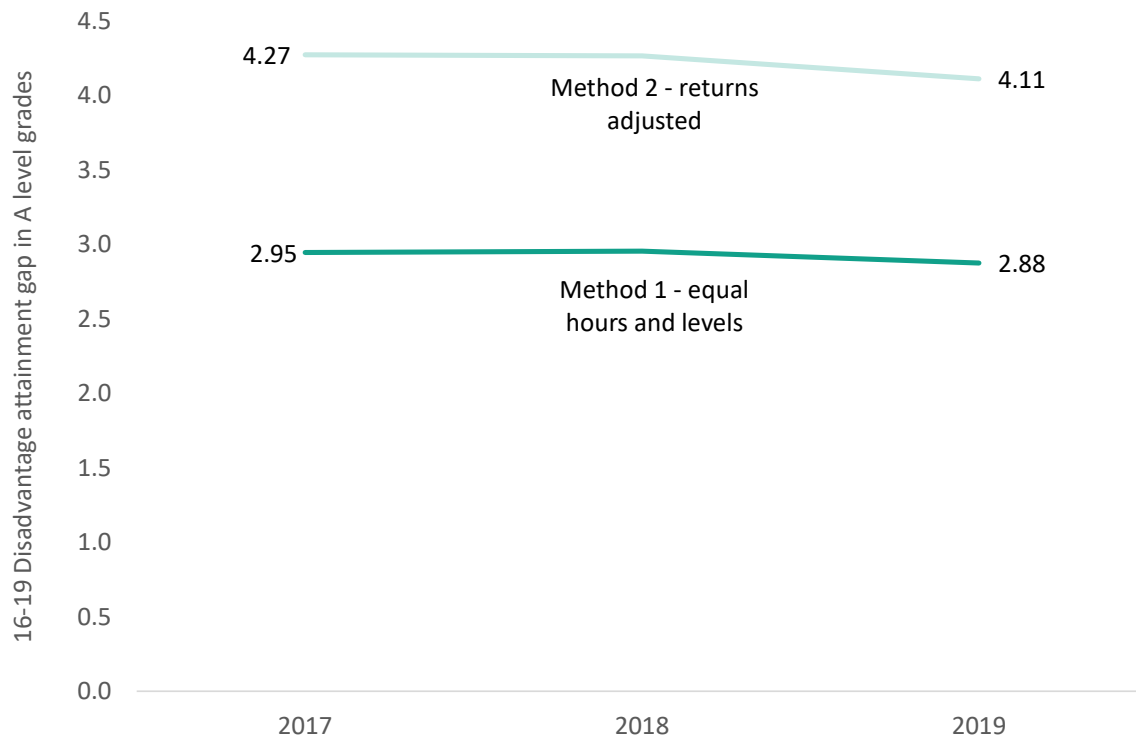
Figure 1 demonstrates that under either approach, there is a clear disadvantage attainment gap that persists into 16-19 education, with non-disadvantaged students around three A level grades behind disadvantaged students under the equal hours and levels methodology, and over four A level grades behind under the returns adjusted methodology. There have been minor variations since 2017, but both measures have remained broadly stable.

The overall gap appears wider under method 2 than it does under method 1. We would expect the disadvantage attainment gap to be wider under method 2, as non-disadvantaged students enter more academic qualifications on average, as demonstrated in figure 2. If we increase the value of these relative to everything else, we therefore increase the gap.

Figure 2: Average number of entries into each qualification type for disadvantaged and non-disadvantaged students
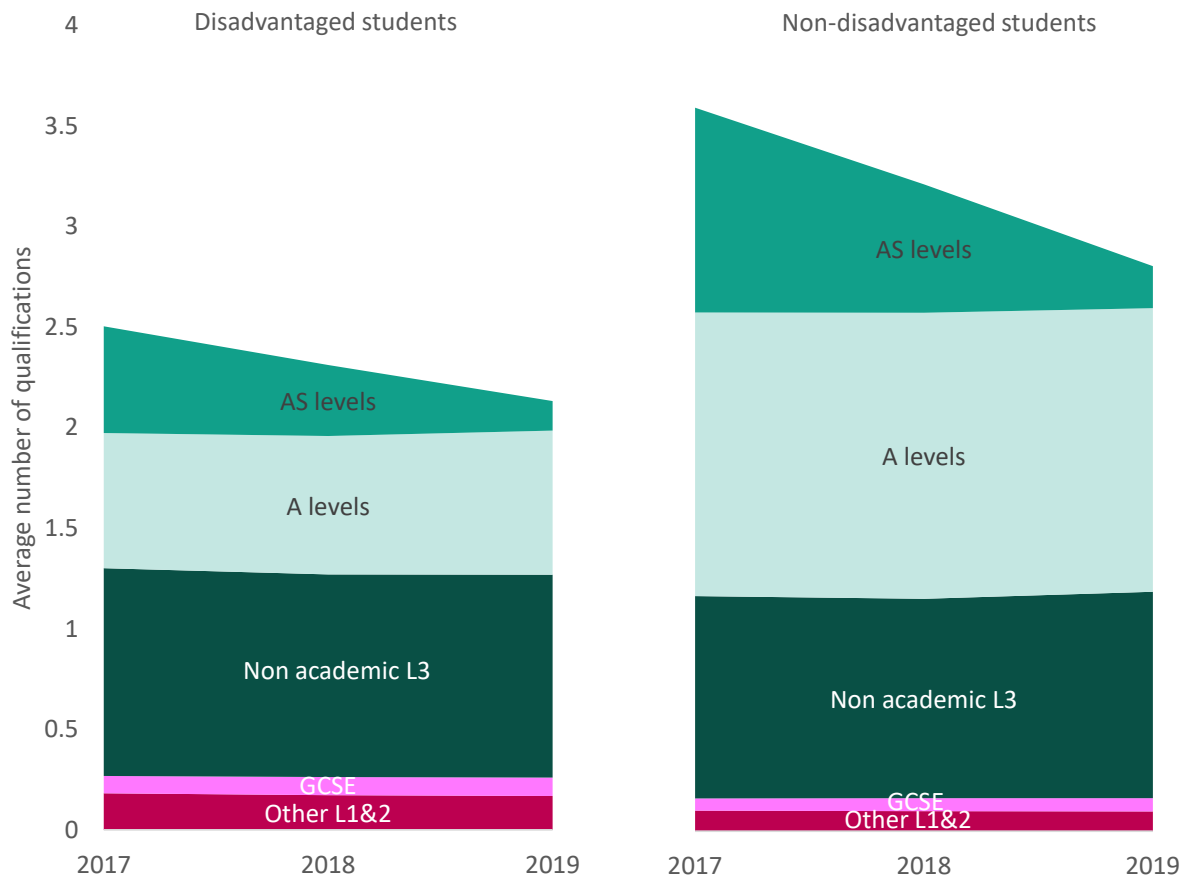
Figure 2 shows that disadvantaged students on average enter fewer qualifications than their non-disadvantaged peers.

As disadvantaged students are less likely to hold more than 3 qualifications, it follows that the volume of qualifications entered will contribute to the size of the overall disadvantage attainment gap, which is based on student's best three results.

The other key trend we see from figure 2 is a decrease in entries to AS levels since 2017. Although the trend is present for all students, it is most prominent for non-disadvantaged students, who were previously more likely to enter AS levels, but have seen entries drop at a greater rate.

This fall in entries is likely to be a combination of two factors. Firstly, the introduction of the decoupling policy, whereby reformed AS levels no longer count towards a full A level. The first reformed AS levels were taught from 2015/16, meaning the first students to hold them will be those finishing a two-year study programme in 2016/17. Further reformed AS levels have been introduced gradually in each year since then, replacing the legacy versions that counted towards full non-reformed A levels. Secondly, the period 2017 to 2019 was one of reductions in real-term funding per student. Between 2012 and 2019 funding per 16-18-year-old student fell by 16 per cent. This may also have played a part in the reduction of provision. Previous EPI research has shown a weak but discernible relationship between changes in funding and reductions in provision at an institution level (Dominguez-Reig and Robinson 2019).

Although this change in entry patterns is substantial, it will be mitigated by the fact that many students previously taking AS levels commonly entered another three full qualifications. As our measure only includes students' best three qualifications, it is not particularly sensitive to changes for groups taking more than this number.

Furthermore, figure 2 shows total qualification entries without any discounting (the process whereby qualifications with significant overlap in content are not allowed to count twice). In the overall measure, where discounting is applied, students that entered AS levels that also went onto take the full A level, would only have the A level result counting in their best three point score, as the AS level would be discounted. This means that although the change in AS level entry patterns is notable, the impact that feeds through to the overall disadvantage attainment gap measures will be less pronounced.

However, the greater narrowing of the disadvantage attainment gap seen under method 2, may be explained by the fact that the number of entries to academic subjects (driven by AS levels) is decreasing at a faster rate for non-disadvantaged students. The greater weight given to academic qualifications under method 2 will magnify this effect.

Table 2: Percentage of students at the end of 16-19 study with at least one entry by qualification type[2]

Qualification type		2017	2018	2019
Academic level 3		55%	53%	51%
Of which:	A level	46%	46%	46%
	AS level	52%	42%	17%
Non-academic level 3		37%	37%	38%
GCSE		19%	20%	20%
Other level 1&2 qualifications		34%	33%	32%

Table 2 shows the percentage of students with one or more entries to the qualification types listed. The figures sum to more than 100 per cent as students that entered a combination of qualifications, such as A levels and BTECs, will appear in more than one qualification type. A levels clearly dominate, and will have a corresponding impact on the overall gap measure.

In addition to the number of entries to each qualification type, the relative attainment of non-disadvantaged and disadvantaged students within them, will affect the overall measures. This is examined further in figure 3 which follows.

---

[2] GCSE does not include English or Maths resits unless the student made positive progress since the end of key stage 4.

Figure 3: Disadvantage attainment gap per qualification by type, 2017 to 2019
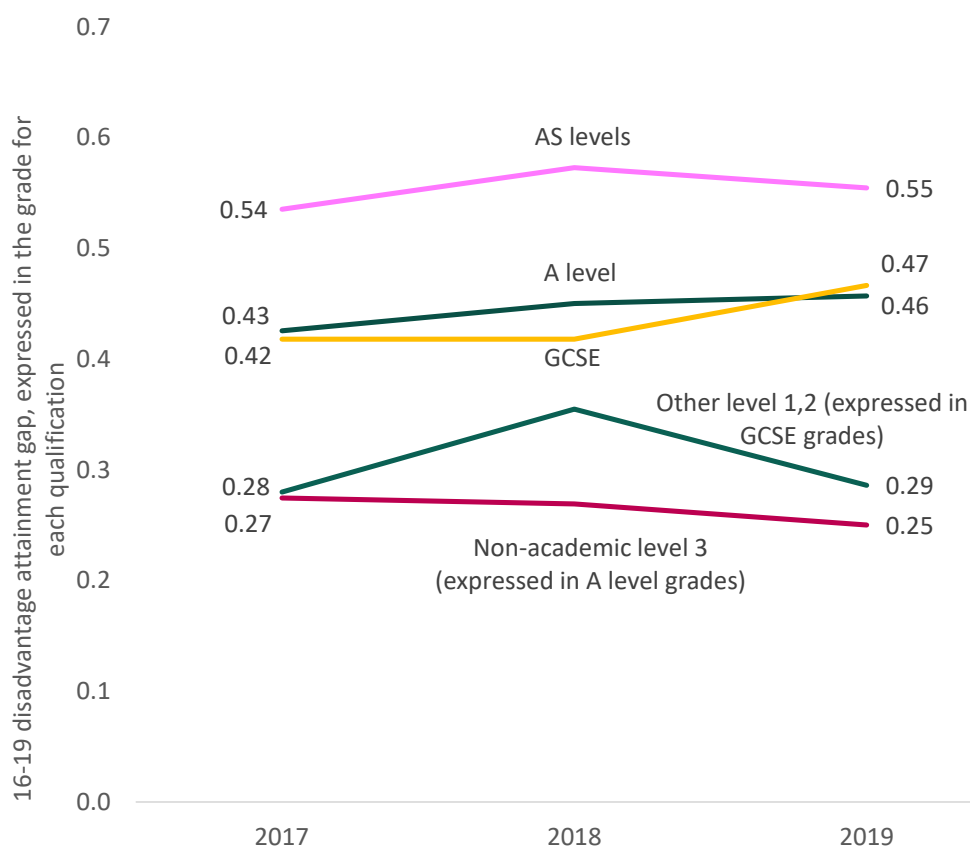
Figure 3 shows how the 16-19 disadvantage attainment gap varies by qualification type.

Students are now only included in these measures if they have entered at least one eligible qualification, and can be included in more than one measure if they entered a mixed programme of qualifications. For example, a student that entered a GCSE and an A level would be included in both the GCSE and A level measure for the purpose of this chart.

The lines represent the disadvantage attainment gap, expressed in terms of the relevant qualification grades. For example, in 2019, disadvantaged students that had entered A levels, on average achieved half a grade less than non-disadvantaged students per entry.

Disadvantaged students also achieved half a GCSE grade less per entry than non-disadvantaged students in 2019. However, GCSEs are lower-level qualifications than A levels and also have far fewer teaching hours; a grade difference at GCSE level therefore corresponds to far less than a grades difference in an A level. This combined with the fact that A levels are far more commonly entered during this phase, means that the relative contribution of GCSEs to the overall measure will be small.

The trend for A levels and AS levels most closely represents the overall trend shown in figure 1. Although less than half of students enter A levels, they are still more commonly taken at 16-19 than any other qualification type, so we would expect them to be a key driver of the overall trend. It should be noted that although the disadvantage attainment gap for AS levels appears wider than for A levels when expressed in terms of grades, the interval between AS level grades is half of that of A levels.

The disadvantage attainment gap for non-academic level 3 qualifications per entry, is substantially narrower than that for A levels.

Figure 4: 16-19 attainment, distribution of point scores by disadvantage status, method 1, 2019[3]

Figure 4 shows the distribution of point scores at 16-19 by disadvantage status in 2019. The low end of the distribution appears similar between disadvantaged and non-disadvantaged students. However, for disadvantaged students there is no peak in the middle of the distribution, as seen for non-disadvantaged students. Instead, we see a constant decrease in the number of disadvantaged students achieving increasingly higher total point scores. Indeed, whilst the modal value for non-disadvantaged students is 110 points, broadly equivalent to a B and 2 C's at A level, the modal value for disadvantaged students is 0, with 17 per cent of disadvantaged students having this score. These cases represent students that have only entered qualifications below level 1 during their 16-19 study, or did not achieve a pass grade in any of the qualifications they did enter.

The mean and median for both disadvantaged and non-disadvantaged students are marked on the chart. In this instance the (horizontal) gap between the median point scores, is greater than the (horizontal) gap between the mean point scores between disadvantaged and non-disadvantaged students.

---

[3] A difference of 10 points can broadly be interpreted as equivalent to achieving a total of one grade more in a qualification of the same level and with teaching hours similar to an A level.

Figure 5: Average key stage 4 English and maths prior attainment
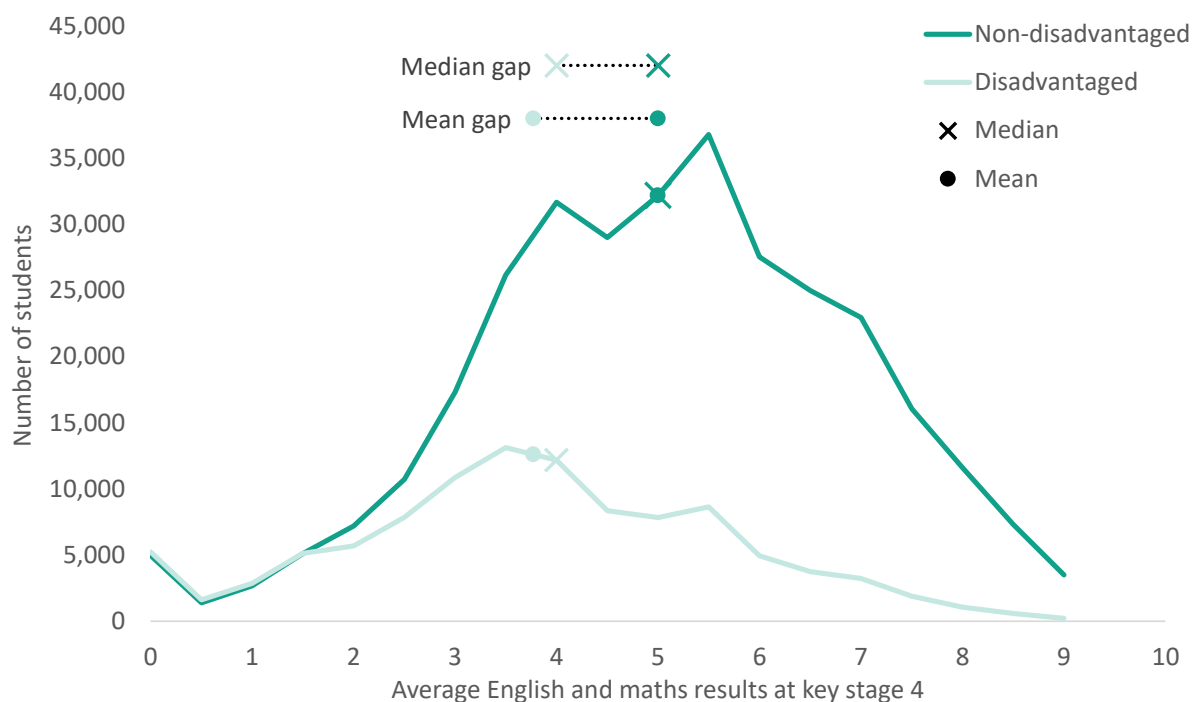
For comparison, figure 5 shows the distribution in key stage 4 prior attainment for the same cohort of students. Unlike the 16-19 attainment distributions, the key stage 4 distribution is more normal, reflecting the more standard programme of study at this key stage, along with higher completion rates.

In the key stage 4 distribution, the difference between the mean point score of disadvantaged and non-disadvantaged students is slightly greater than the difference between the median difference.

Given the lack of normality in the distribution, we also tested the difference in attainment based on the difference in mean and median percentile rank between disadvantaged and non-disadvantaged students.

Table 3 below shows the resulting disadvantage attainment gap for students at the end of 16-19 study under different methodologies. It also shows the same metrics based on the key stage 4 results of these same young people, when they were at the end of secondary school.

Table 3: Disadvantage attainment gap under different methodologies for students at the end of 16-19 study in 2019, and when these same students were at the end of key stage 4

Phase	Difference in mean rank	Difference in median rank	Difference in mean standardised point score	Difference in median standardised point score
End of 16-19 study	0.17	0.26	0.60	0.93
Key stage 4 (same cohort)	0.18	0.21	0.64	0.52

Table 3 demonstrates that the different attainment distributions for key stage 4 and 16-19 lead to different interpretations, dependent upon the measure used. Median based attainment gap

measures suggest a bigger gap in 16-19 education, while mean based gap measures suggest a bigger gap at key stage 4.

The difference in the shape of the key stage 4 and 16-19 distributions means that simplistic comparisons using aggregate statistics, such as mean rank or point score differences, should be treated with caution.

16-19-year-olds without qualifications

Not all students continue from key stage 4 to undertake qualifications during the 16-19 phase. Around four per cent of 16 year olds transition into apprenticeships, three per cent into employment and five per cent have no recorded employment, education or training destination. [4]

Though our headline measures consider the disadvantage attainment gap for those in 16-19 education, we have also produced a supplementary measure that includes students that did not continue education and complete a study programme in the 16-19 phase. The supplementary measure has been created by identifying students that appeared in state funded key stage 4 data but had not appeared in the 16-19 education datasets by age 19. For the purpose of this measure they have been allocated zero points, reflecting that they had not entered any qualifications in this phase.

In addition, students on an apprenticeship programme have been included and allocated points equivalent to the average grade in Tech-levels (vocationally oriented level three qualifications, rescaled to be equivalent in size to two A levels) for level 3 apprenticeship programmes, and the average grade in a level 2 technical certificate (rescaled to be equivalent in size to 5 GCSEs) for level 2 apprenticeship programmes.

---

[4] https://explore-education-statistics.service.gov.uk/find-statistics/key-stage-4-destination-measures/2018-19

Figure 6: Supplementary measure, 16-19 disadvantage attainment gap in A level grades - including apprentices and those that did not continue education post 16[5]
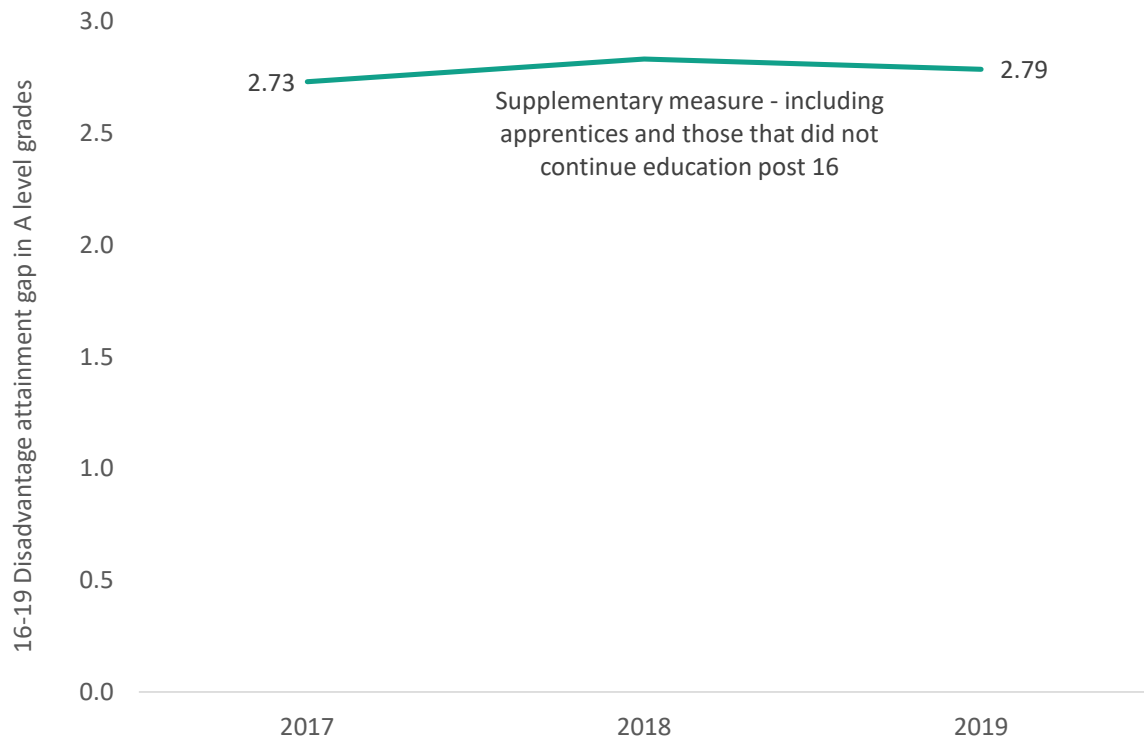
Figure 6 shows a slightly narrower disadvantage attainment gap than the previous measures, of around 2.8 A level grades in 2019. This is as a result of including apprentices in the measure. Although non-disadvantaged students are slightly more likely to complete a higher level apprenticeship, the resulting point score gap for these students is smaller than for those included in the main attainment measures shown in figure 1.

The effect of including students that did not continue in education post 16 is to slightly widen the disadvantage attainment gap, as they have been allocated 0 points, and disadvantaged students are over-represented within this group. However, when combined with the effect of including apprentices, the overall effect is that the supplementary measure appears narrower. We have produced this measure to demonstrate the effect of including the entire cohort of 16-19 year-olds for completeness. However, it is not used in the more detailed analysis that follows to not conflate differences in attainment with differences in continued participation in the education system, or likelihood to take up an apprenticeship.

---

[5] Note that there is a larger number of 19 year olds that did not continue education post-16 included in the figure for 2017, as these students would have finished key stage 4 in 2014, prior to the compulsory education age being raised to 18

Variation by student and institution characteristics

In this section, we explore the difference in 16-19 attainment outcomes by student and institution characteristics. The charts presented compare the average total point score under method 1 within different groups, rather than the difference between groups expressed in equivalent A level grades, as in the previous section. This is so the absolute level of attainment, not just the relative differences between groups can be examined. Throughout, a difference of 10 points in the attainment measure equates to one A level grade or the difference between a Pass and a Merit in a level 3 BTEC national extended certificate (or other level 3 qualification with similar teaching hours). We later go on to examine how these findings vary when holding a range of other characteristics constant, for example prior attainment.

Figure 7: 16-19 attainment by gender and disadvantage status
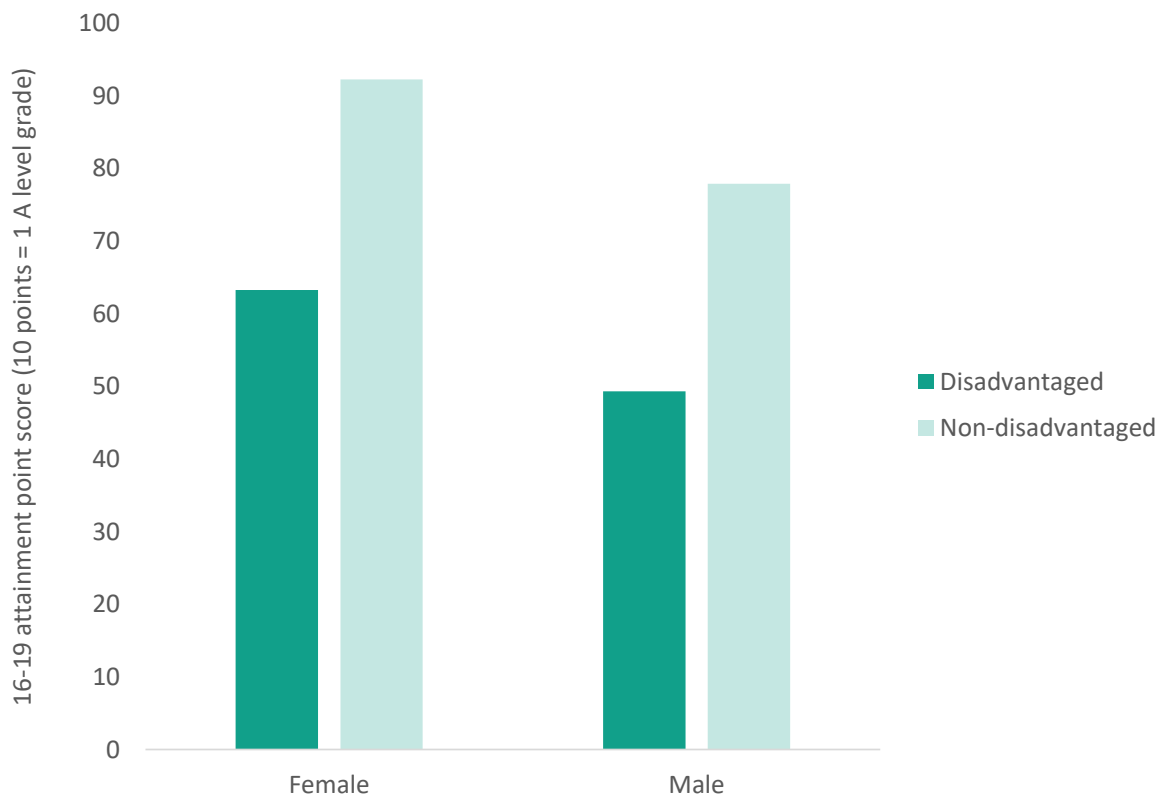
Figure 7 shows 16-19 attainment, split by gender and disadvantage status.

We see that overall, for both disadvantaged students and their peers, female students are likely to have a higher point score than males.

Figure 8: 16-19 attainment by first language and disadvantage status
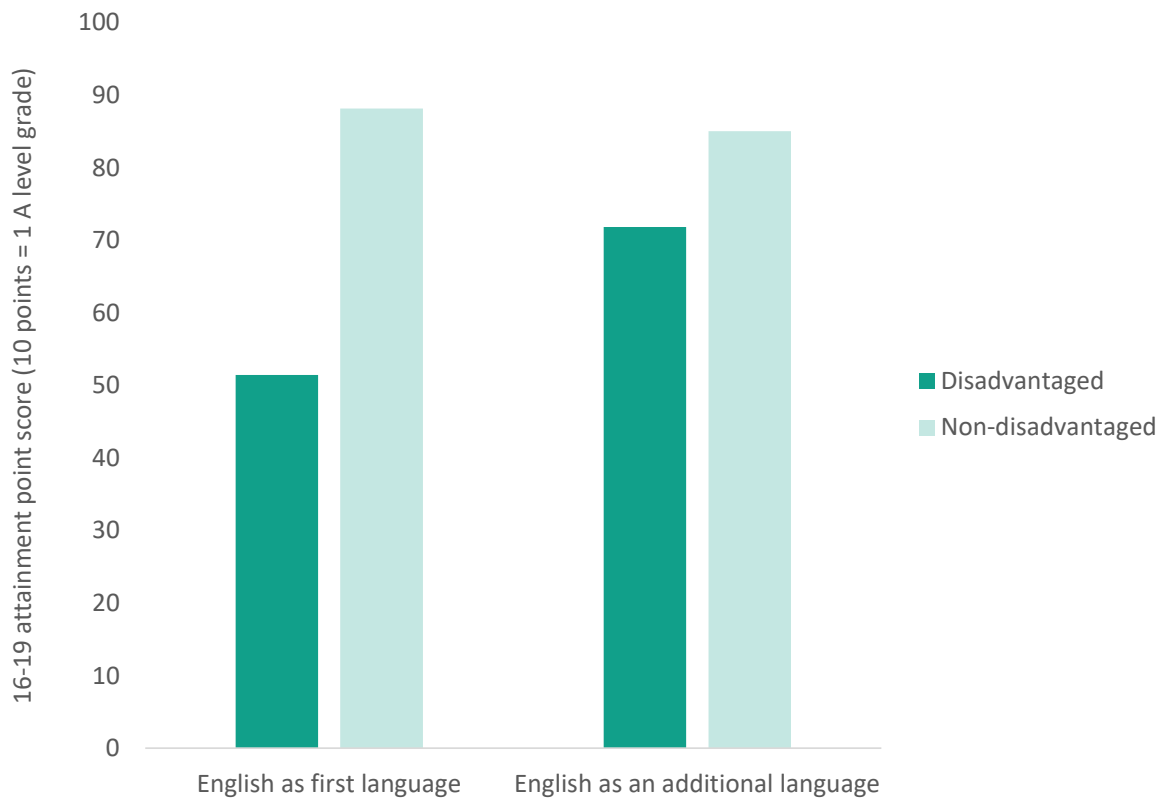
Figure 8 shows the average 16-19 attainment for students by whether or not English was their first language (EAL), and disadvantage status.

The attainment of non-disadvantaged students appears similar regardless of students' first language. However, disadvantaged students with EAL appear to have much higher 16-19 attainment that students whose first language is English, resulting in a considerably smaller disadvantage attainment gap for EAL students. This is a finding we revisit in section 5.

Figure 9: 16-19 attainment by special educational needs status and disadvantage status
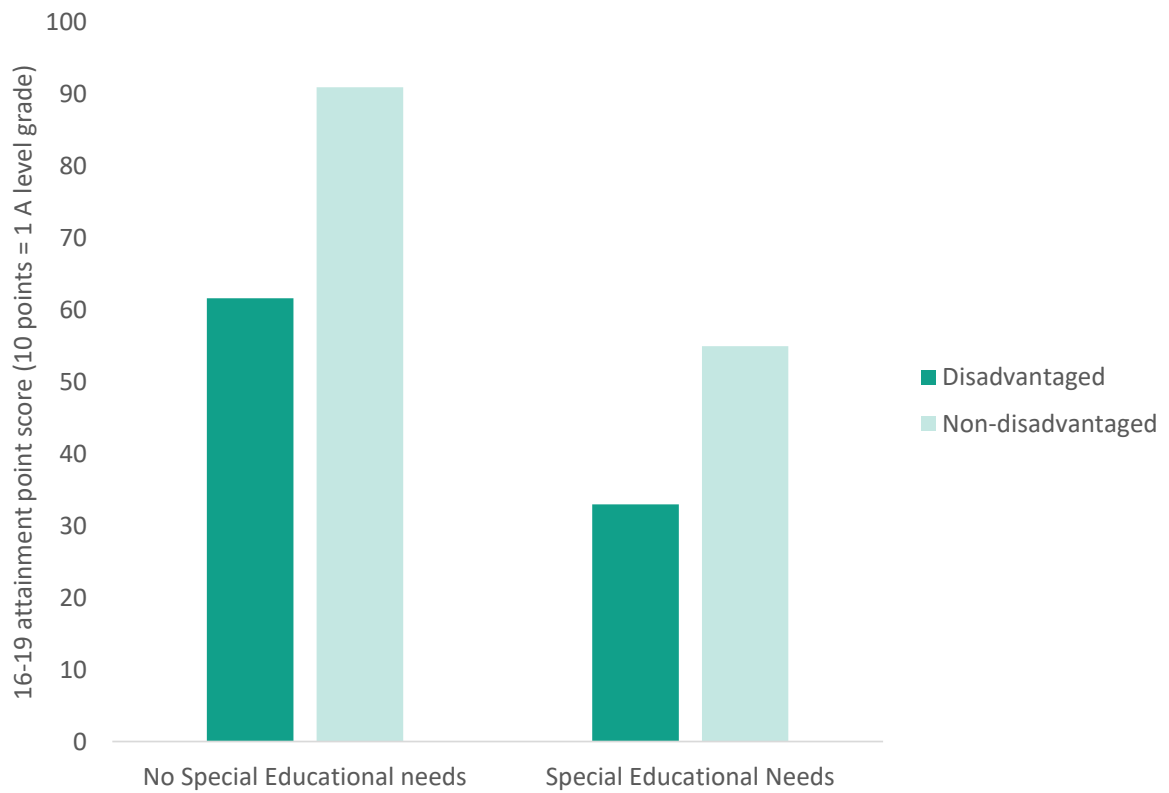
Figure 9 shows that those with special educational needs, on average have lower 16-19 attainment, and that disadvantaged students are likely to have lower attainment than their non-disadvantaged peers, regardless of whether they have special educational needs.

Figure 10: 16-19 attainment by ethnic background and disadvantage status[6]
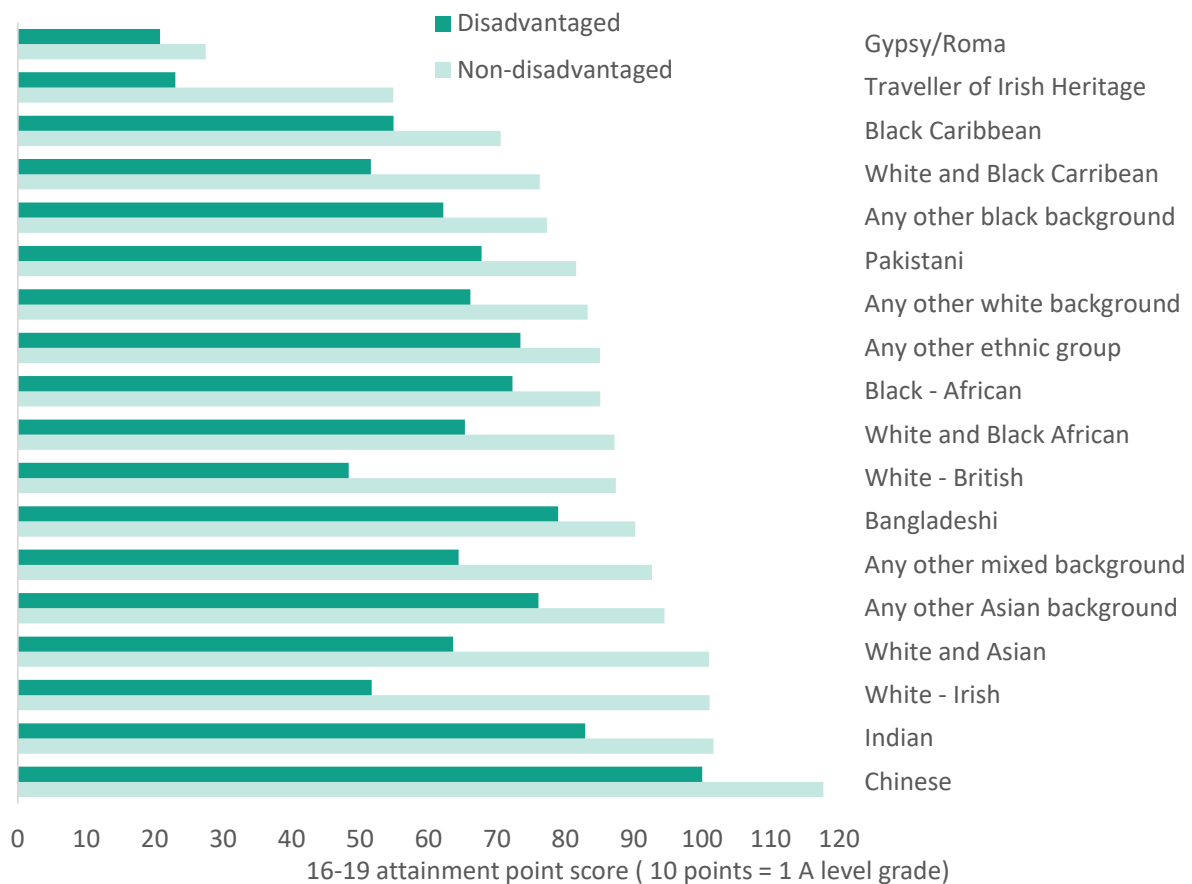
Figure 10 shows the total point score achieved by disadvantaged (non-disadvantaged) students from different ethnic backgrounds under method 1 – the equal hours and levels methodology. Only disadvantaged students from a Gypsy/Roma background or travellers of Irish heritage had lower attainment than disadvantaged White British students. The results of non-disadvantaged students are higher than those of disadvantaged students across all ethnicities, but poverty appears to matter less amongst some groups. The disadvantage attainment gap is largest for White – Irish and White – British students, followed by White and Asian students.

The trends are similar to what we see at key stage 4, where disadvantaged White British students achieve amongst the lowest of all disadvantaged students(Department for Education 'Key Stage 4 Performance 2019 (Revised)').

---

[6] Note that the number of travellers of Irish heritage students is small relative to the other categories, so the mean point score will be more sensitive to individual results. Students that could not be linked back to a school census record have been excluded from this analysis. A difference of 10 points can be interpreted as a grades difference in an A level, and broadly similar to the difference between a pass and a merit, or a merit and a distinction in a BTEC with similar teaching hours.

Figure 11: 16-19 attainment by school or college type and disadvantage status[7]
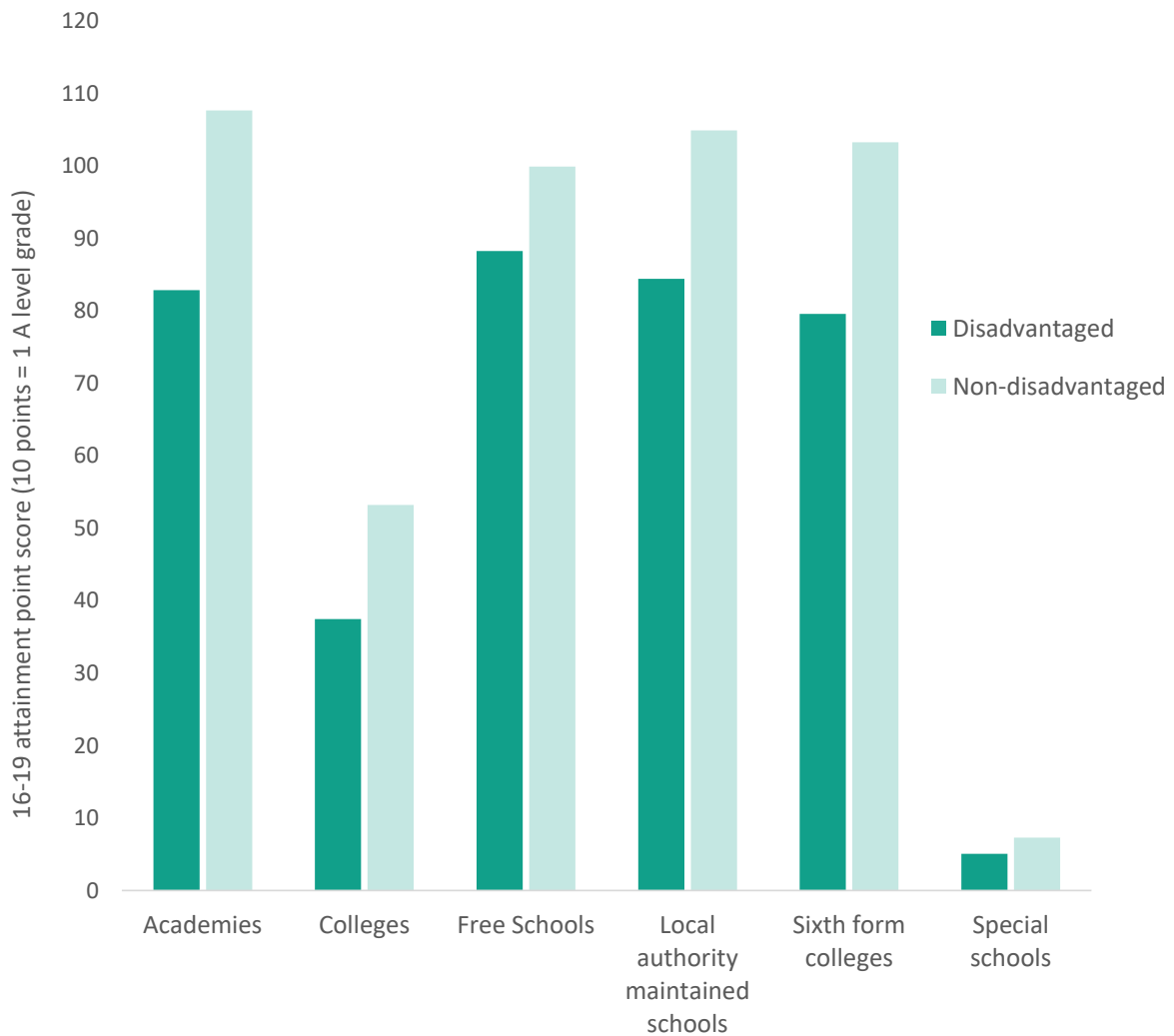
Figure 11 looks at the average total point score within school and college types. The chart shows that across all school and college types, non-disadvantaged students achieve higher results than disadvantaged students. The differences seen between schools, sixth form colleges and other FE colleges will reflect a combination of the differences in grades achieved, and the number and size of qualifications entered. Students attending Further Education colleges achieve a lower total point score than students in other forms of mainstream provision (i.e. those within academies or local authority-maintained schools). Most, though not all, of this difference can be explained by the lower prior attainment of students attending further education colleges. We explore this in greater detail in section 5.

---

[7] As with the previous chart, a difference of +10 can be interpreted as roughly a grades difference in an academic or vocational qualification with similar teaching hours to an A level.

4 - Local authority analysis

The tables below show the ten local authorities that have the widest and narrowest 16-19 disadvantage attainment gap, using method 1. Full results for all local authorities can be found in Annex B and are displayed for England and London in map 1.

At a national level, the disadvantage attainment gap is measured as the equivalent number of A level grades that disadvantaged students are behind non-disadvantaged students. At local authority level, we have looked at the number of A level grades that disadvantaged students within each local authority, are behind non-disadvantaged students nationally. Using a national, non-disadvantaged comparator is preferable as it allows for meaningful comparisons of how well disadvantaged children are achieving between local authorities. If the disadvantage attainment gap were measured entirely within local authorities, some may have a very small gap, purely as a result of non-disadvantaged students having low attainment.

Table 4: The 10 local authorities with the widest disadvantage attainment gap, 2019

Local authority	Disadvantage attainment gap, A level grades
Knowsley	5.4
North Somerset	4.8
Stockton-on-Tees	4.7
Torbay	4.4
Swindon	4.4
Derby	4.3
Barnsley	4.3
Hartlepool	4.2
South Gloucestershire	4.1
West Sussex	4.1

Table 5: The 10 local authorities with the narrowest disadvantage attainment gap, 2019

Local authority	Disadvantage attainment gap, A level grades
Southwark	-1.2
Redbridge	-0.5
Ealing	-0.5
Sutton	-0.2
Merton	-0.2
Newham	-0.2
Islington	-0.1
Bexley	0.0
Wokingham	0.1
Harrow	0.1

Table 5 shows that the local authorities with the smallest disadvantage attainment gap are mostly in London. Where values are negative, this implies that disadvantaged students in these areas, on average, are ahead of non-disadvantaged students nationally in terms of the grades they achieved.

The local authorities with the widest disadvantage attainment gap are spread more evenly across England, with no clear geographic pattern. However, there is correlation with the trends seen at key stage 4. The Education Policy Institute's 2020 annual report (Hutchinson, Reader, and Akhal 2020) shows that based on English and maths GCSE results, Knowsley, Torbay, Derby and South Gloucestershire also appear in the 10 local authorities with the widest gap at key stage 4. Likewise, Redbridge, Ealing and Newham all still appear in the 10 local authorities with the narrowest gap at key stage 4.

This finding is as we would expect given that prior attainment explains a large proportion of the variation we see in 16-19 attainment outcomes, as discussed in the sections which follow.

Map 1: 16-19 disadvantage attainment gap by Local Authority in equivalent A level grades, England and London, 2019[8]
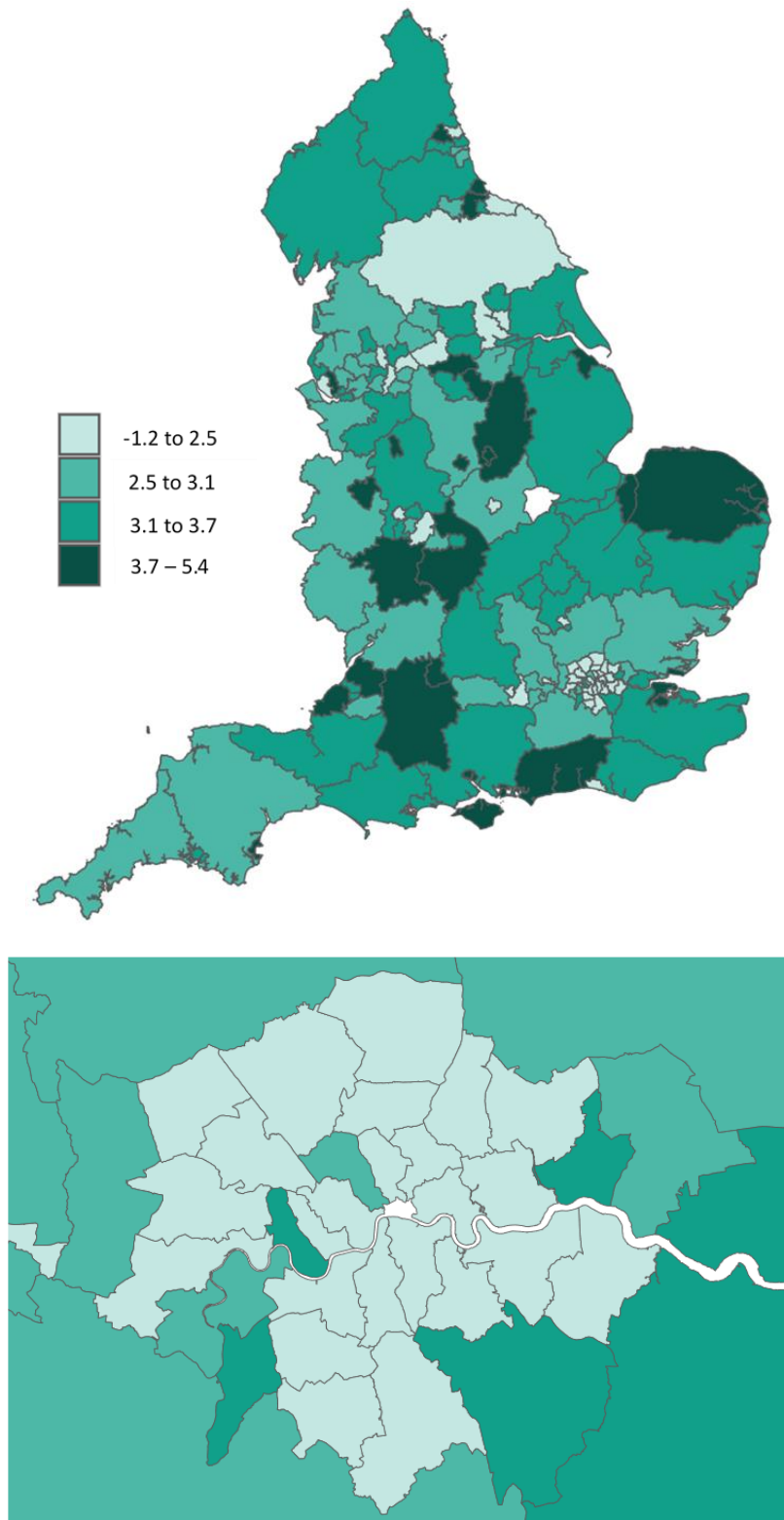
Legend:
- -1.2 to 2.5
- 2.5 to 3.1
- 3.1 to 3.7
- 3.7 – 5.4

5 - Factors associated with 16-19 attainment

In this section we provide a deeper understanding of the differences in attainment across the different student and provider characteristics observed in the previous chapter. To do so we have developed a series of regression models to investigate the association of different characteristics with 16-19 attainment, whilst holding other factors constant.

All analysis presented in this section of the report is based upon the method 1 version of the 16-19 attainment measure described in section 2, for students finishing 16-19 study in the 2018/19 academic year. Our analysis showed that results based on the method 2 methodology are similar. The main difference is that under method 2, the disadvantage attainment gap appears wider, but more of it can be explained in terms of prior attainment. This is because academic qualifications have a stronger association with English and maths results at key stage 4.

It should be noted that whilst this analysis increases our understanding of the factors most strongly associated with the disadvantage attainment gap, our research design does not allow us to attribute causality to any particular factor.

A brief description of the model types applied can be found in the methodology section of this report. The key findings from these models are discussed in this section, but full results from all models can be found in Annex D.

Demographic characteristics

Figure 12 below looks at how the relationship between demographic characteristics and attainment changes as other factors are controlled for.

Figure 12: regression coefficients for demographic characteristics when controlling for other factors[9]

Regression coefficients (10 points equates to roughly one grades difference in a qualification with similar teaching hours to an A level)
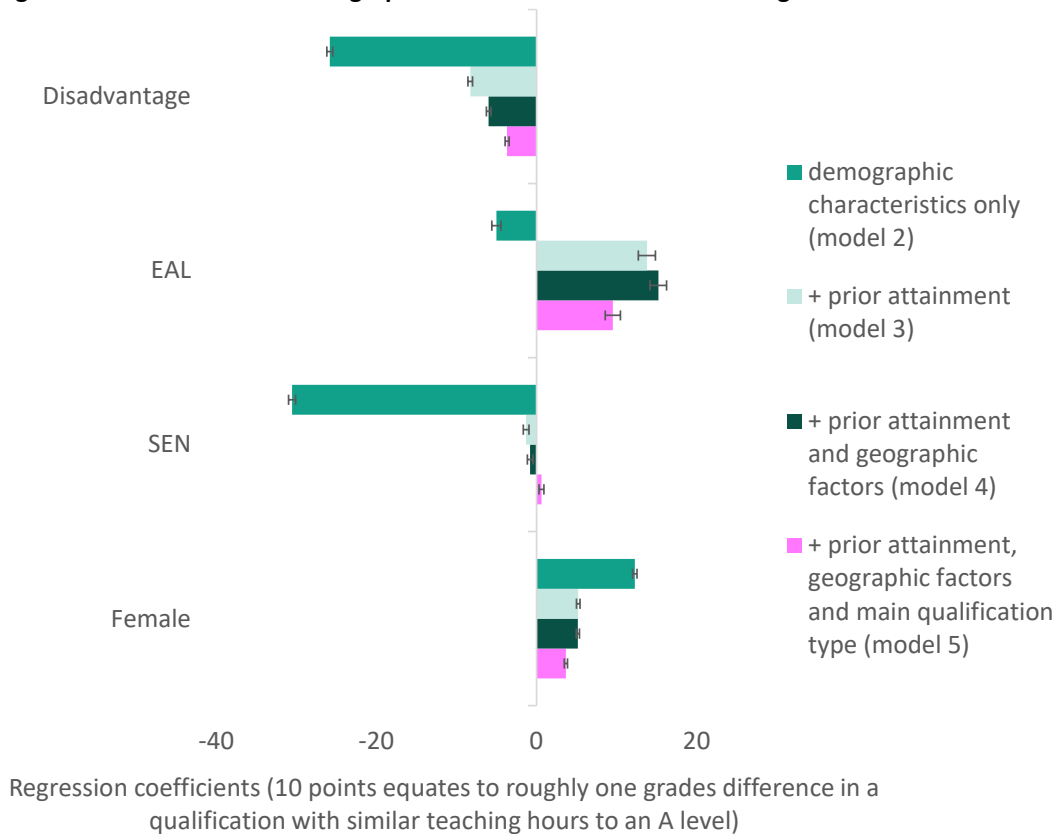
We see that the influence of disadvantage (in terms of the regression coefficients) on 16-19 attainment decreases with every additional set of variables we control for. In particular, we see a sizeable decrease when prior attainment is added, suggesting that prior attainment may be a significant driver in the disadvantage attainment gap. There is more detail on this in the following chapter.

The influence of special educational needs (SEN) on 16-19 attainment disappears once prior attainment is controlled for, indicating that students with SEN, who tend to have lower key stage 4 attainment, do not fall further behind during the 16-19 phase, on average.

Unlike other basic demographics, the influence of English as an additional language (EAL) reverses once prior attainment is controlled for. This suggests that whilst the average student with EAL has lower prior attainment than average, their 16-19 attainment is higher than students with similar prior attainment, i.e. they make good progress from a low baseline. This may be because prior

---

[9] A coefficient of -10 can be interpreted as meaning that on average, disadvantaged status is associated with achieving one grade less in a level 3 qualification with similar teaching hours to an A level, than students that are similar in terms of other characteristics. The error bars displayed on the charts throughout this section represent the 95 per cent confidence interval of the associated regression coefficient i.e. whether the attainment for the characteristic shown is statistically significantly different from the reference characteristic e.g. female students have statistically significantly different results from males.

attainment underestimates ability for students with English as an additional language. That is, students may have trouble being taught in their second language, which diminishes through time as their fluency improves.

In a direct comparison (seen previously in figure 7) female students have 16-19 attainment one grade higher than male students on average. However, this difference more than halves once the higher key stage 4 prior attainment is considered.

Figure 13: regression coefficients by ethnicity, relative to White British students, before and after including prior attainment

Regression coefficients (10 points equates to roughly one grades difference in a qualification with similar teaching hours to an A level)
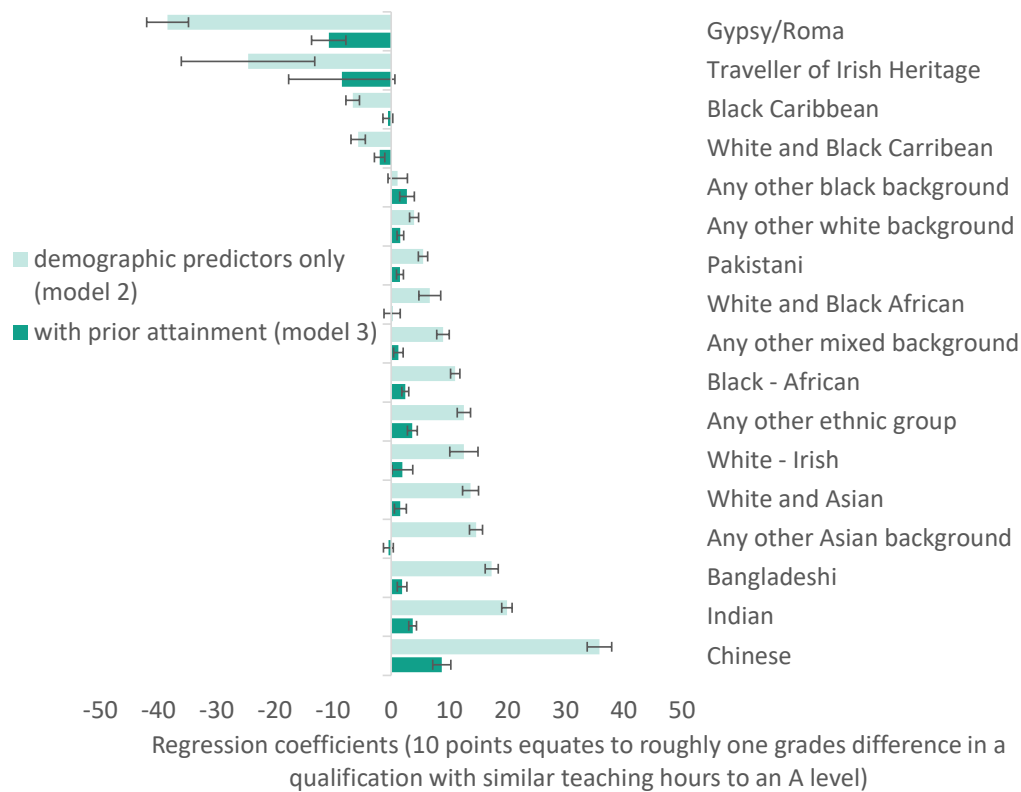
The coefficients presented in figure 13 are relative to White British students. For example, having controlled for demographic characteristics and prior attainment, Chinese students are likely to achieve around a grade higher than White British students.

This figure tells a slightly different story to figure 10, which looked at attainment by ethnicity and disadvantage without controlling for other factors. In figure 10, disadvantaged students from nearly all ethnic backgrounds achieved a higher outcome than disadvantaged White British students, but this was not the case for non-disadvantaged students.

Figure 13 shows that when we control for other demographic factors, the coefficients associated with most ethnic backgrounds is positive, indicating an association with higher attainment than White British students. Once prior attainment is controlled for the coefficients reduce substantially, in most cases to less than a third of an A level grade (or equivalent). As with the previous charts, the fact that the coefficients reduce so greatly when prior attainment is controlled for, does not diminish the relevance of ethnicity. It rather suggests that the association between ethnicity and attainment is already there by the end of key stage 4, and that it persists into 16-19 study.

Institution level factors

Figure 14 shows that only eight per cent of the total explained variation in 16-19 attainment can be attributed to differences between institutions. Although the majority is accounted for by within institution differences, this figure is not insubstantial and implies that choice of specific school or college is important and is associated with students' attainment. This figure is also similar in scale to figures observed for key stage 4 (Evans 2007).

Figure 14: Proportion of explained variance that can be attributed to between and within institution differences.

Between institution variance, 8%

Within institution variance, 92%
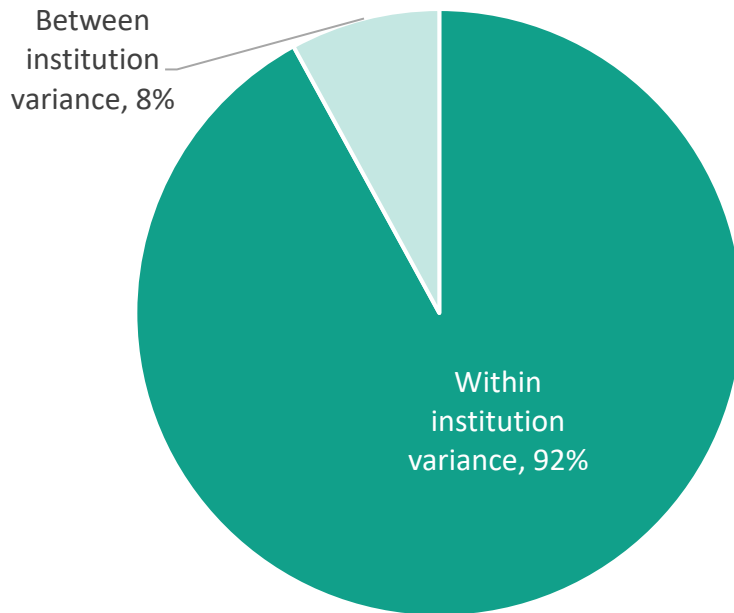
Figure 15 presents the regression coefficients associated with different school or college types relative to local authority-maintained schools.

Figure 15: Regression coefficients by school or college type, relative to local authority maintained, before and after including prior attainment

Regression coefficients relative to LA maintained (10 points equates to roughly one grade difference in a qualification with similar teaching hours to an A level)
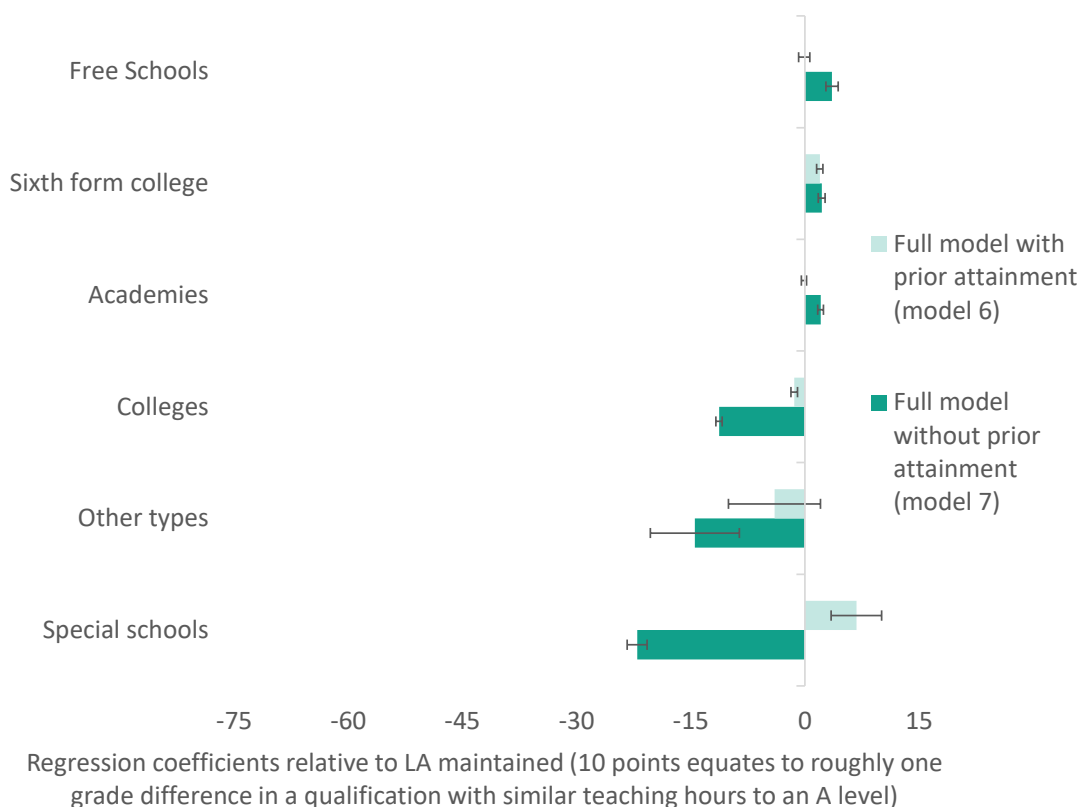
We see that academies, free schools and sixth form colleges have very small coefficients, indicating students achieve similar results to those in local authority schools once other factors have been adjusted for. As such, differences in student characteristics appear to account for the majority of the differences in 16-19 attainment between these institution types seen in figure 11.

Including controls for main qualification type and other student demographic variables, but excluding prior attainment controls, colleges have a negative coefficient of around 11 points. This is equivalent to the difference between achieving a pass and a merit in a BTEC national extended certificate.

Once prior attainment is also controlled for, the coefficient for colleges reduces substantially, to less than 2 points, that is, less than a fifth of an A level grade. This suggests that most of the lower 16-19 attainment observed in colleges is due to differences in student characteristics or types of qualification entered.

Once prior attainment is controlled for, the coefficient for special schools switches from a large negative to a positive score. As we are also including a variable to identify students with Special Educational Needs, these coefficients effectively reflect the attainment of students with SEN in special schools relative to students with SEN in LA maintained school sixth forms.

Figure 16: Proportion of variation in 16-19 attainment explained by differences in prior attainment and institution type
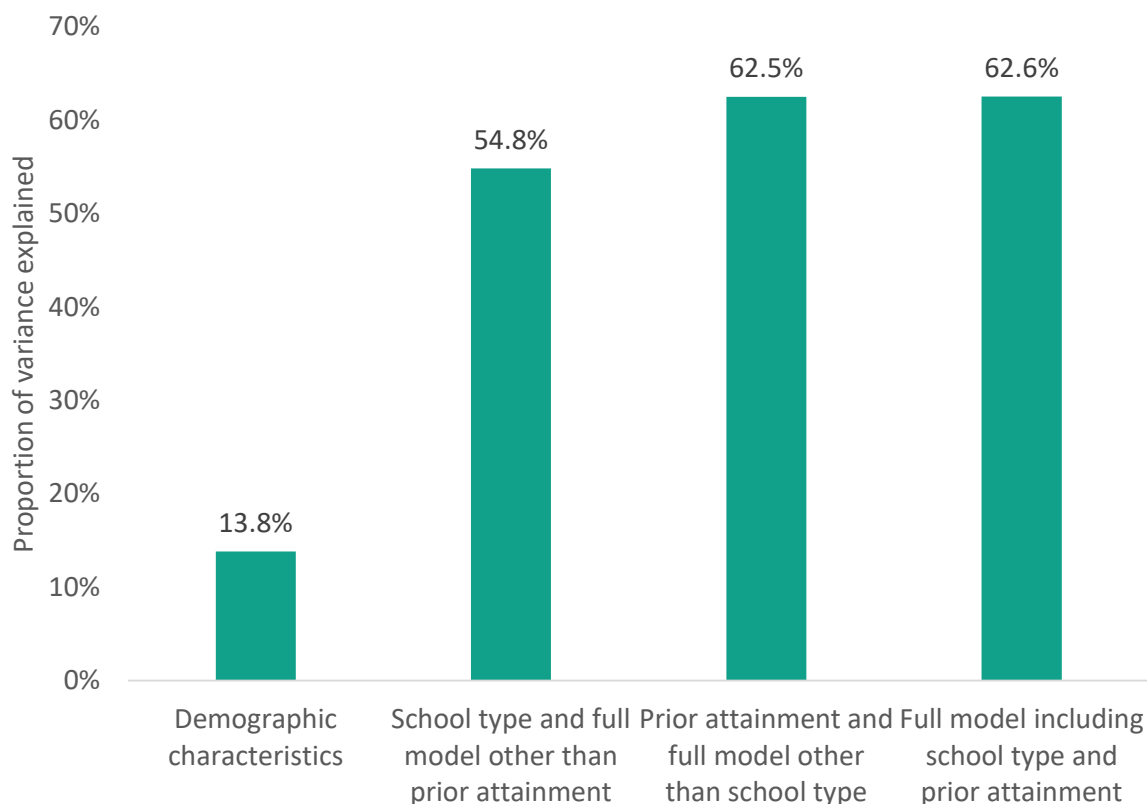
Figure 16 shows the proportion of variation in 16-19 attainment explained by the student and provider characteristics described, in models with and without prior attainment and institution type.

We can see that removing prior attainment from the full model significantly reduced the amount of variation explained, from 62.6 per cent to 54.8 per cent, reinforcing what we know about the importance of prior attainment in explaining differences in 16-19 attainment.

However, removing school type from the full model creates a much more modest reduction in the variation explained, from 62.6 per cent to 62.5 per cent. This suggests that the differences we see in 16-19 attainment between institution types are largely explained by differences in the prior attainment of the students attending them, consolidating the findings emerging from figure 15.

Area and student level disadvantage

The Index of Multiple Deprivation (IMD) gives a value of relative deprivation for every lower layer super output area (LSOA) in the country.[10] It is based on average levels of income deprivation, employment deprivation, health deprivation and disability, education, skills and training deprivation, barriers to housing and services, crime and living environment deprivation.

The disadvantage component of 16-19 funding formula is based on the IMD. For each student living in the 27 per cent most deprived areas institutions receive an uplift. The size of this uplift depends on the level of deprivation (Education and Skills Funding Agency 2019). Here we compare how much

---

[10] LSOAs are geographic areas in England and Wales. They have a minimum population of 1,000 and a mean of 1,500.

of the variation in 16-19 attainment can be explained by student level disadvantage (based on eligibility for free school meals during key stage 4) and the IMD.

Figure 17: Proportion of variation in 16-19 attainment explained by disadvantage status and the Index of Multiple Deprivation[11]
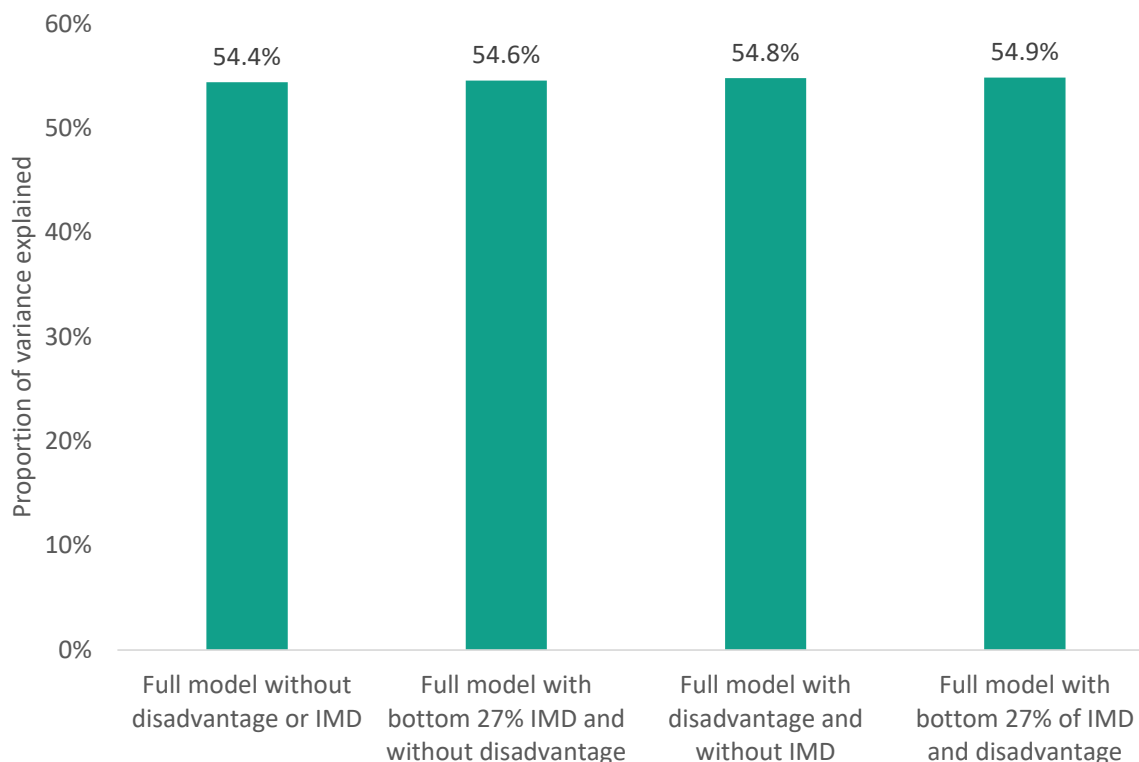
The marginal effect of including student level disadvantage (0.4 per centage points) is greater than that of including IMD (0.2 per centage points). Including both measures improves the explanatory power of the model by 0.5 per centage points. Though these figures appear small, this is in part due to the relatively low proportions of students flagged as either disadvantaged or within the bottom 27 per cent of IMD categories. The full model results in Annex D show that once this version of the IMD measure is included within the model, students flagged as disadvantaged still have a coefficient equivalent to almost three quarters of an A level grade lower than non-disadvantaged students.

Family background and aspirations

In this section we use the Next Steps survey to provide a greater understanding of the influence of parental background and student aspirations. We consider information on parents' education levels

---

[11] To more closely mirror the definitions used in the 16-19 funding model, IMD in the regression models shown in figure 17 has been defined as a percentile rank representing increasing disadvantage for the bottom 27 percent, and 0 otherwise.

Prior attainment in these models has also been adjusted to include binary flags in line with the funding definitions, indicating whether students achieved a grade 4 or above in GCSE (or equivalent) English and/or Maths. This coarser measure of prior attainment explains the lower overall explanatory power of the models shown.

and socioeconomic classification as well as students' plans for continuing education or transitions to employment. As noted in the methodology section, this analysis differs from NPD/ILR based analysis presented previously as it is based on level 3 qualifications only, and from the cohort completing their 16-19 education in 2007.

The main findings are presented in the figures below.

Figure 18: Variation in 16-19 attainment explained by different characteristics (cumulative), 2007
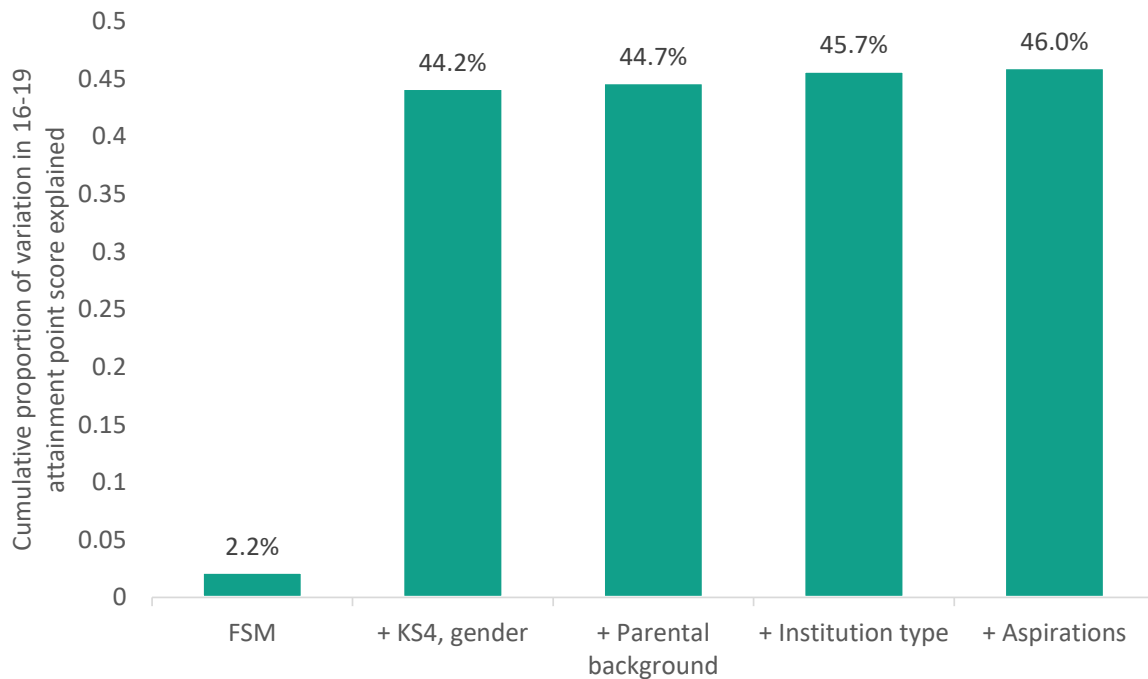
Figure 18 shows that as with the full model on more recent data, much of the variation in 16-19 point scores can be explained in terms of prior attainment. Controlling for gender and free school meal status explains almost 45 per cent of the total variation in 16-19 attainment. However, parental background and aspirations only explain little additional variation (0.5 and 0.3 per cent respectively). Whilst this may seem surprising, it is likely that these factors have already influenced young people's education prior to the 16-19 phase, and have therefore fed into students' key stage 4 results. As such their influence on 16-19 attainment appears diminished once prior attainment is controlled for.

However, this also implies that the absence of these factors in our modelling based on 2019 NPD/ILR data do not represent a significant omission.

Figure 19: Selected coefficients from full 16-19 level 3 average point score regression model, 2007[12]
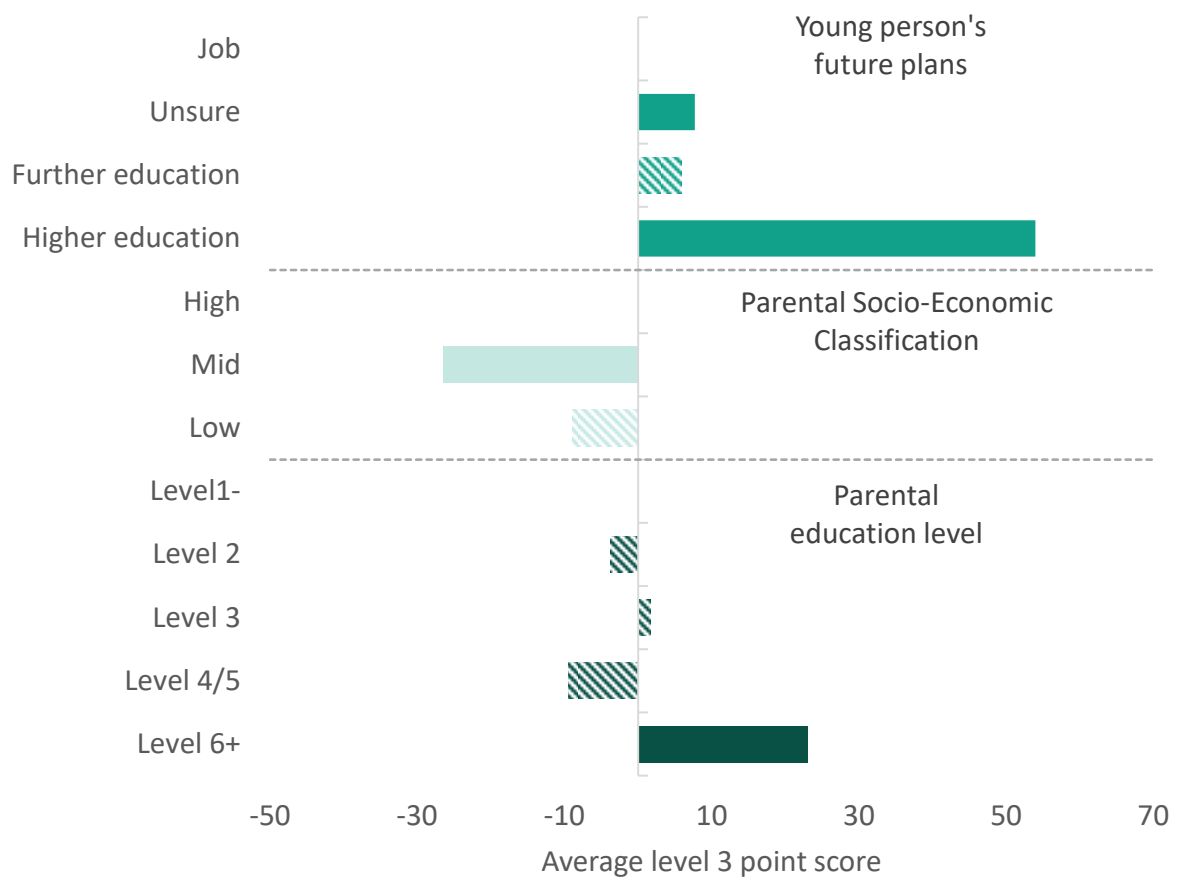
Figure 19 displays the coefficients for the additional variables available in the Next Steps data, from a model also including FSM eligibility, key stage 4 prior attainment, gender and institution type. There is a positive and significant association with 16-19 attainment for students planning to pursue higher education (relative to going straight into the workplace after schooling) and for students with parents holding qualifications of level 6 or above (relative to parents holding qualifications of level 1 or below). There is also a negative coefficient for students with parents from mid-level socio-economic classifications relative to those with parents from high-level socio-economic classifications.

---

[12] Reference categories are displayed on the chart as 0. For example, the coefficient associated with being female is 17, relative to the reference category of being male, displayed as 0. Variables for which the coefficients were not significant in the model are shaded.

6 - Factors associated with the 16-19 disadvantage attainment gap

In this section we more directly consider the factors that may be influencing the gap between the attainment of disadvantaged 16-19 year-olds and their peers.

Figure 20: The coefficient for students' disadvantage status in a simple linear regression model when controlling for an increasing number of variables

Regression coefficients (10 points equates to roughly one grades difference in a qualification with similar teaching hours to an A level)
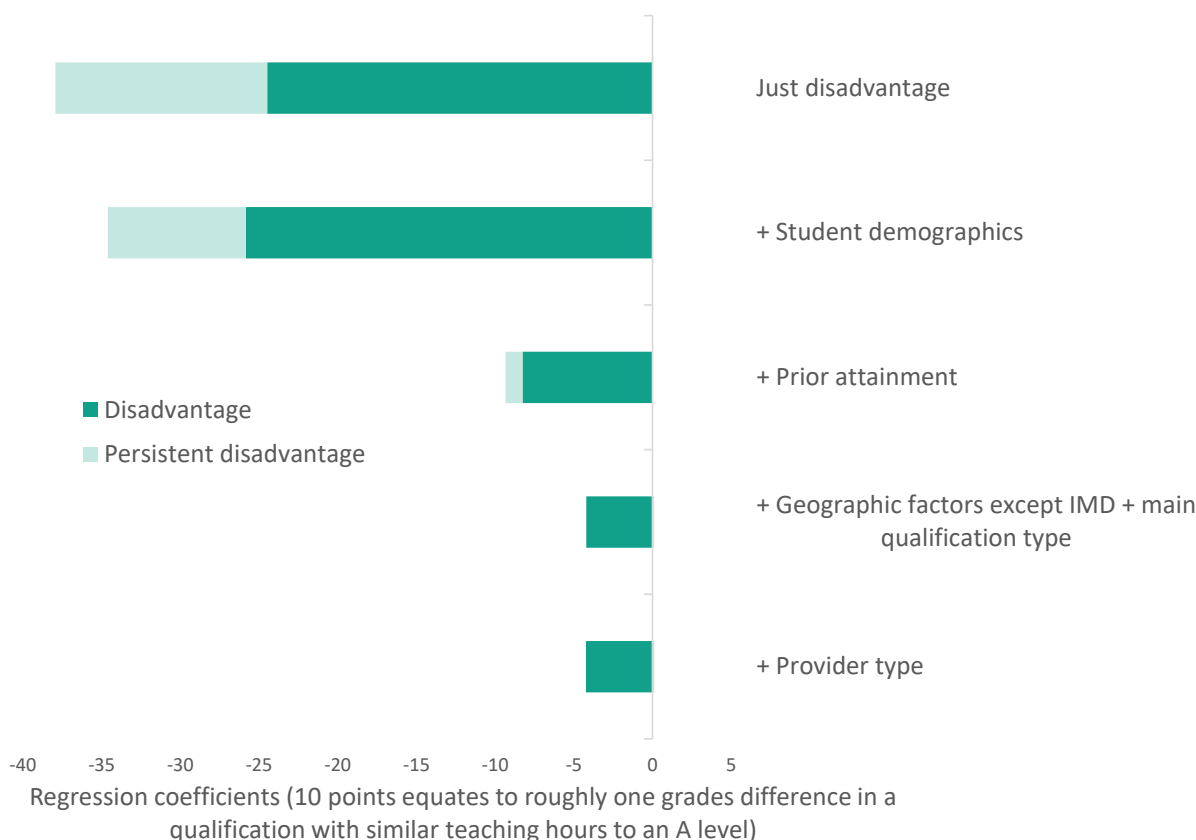
Using similar models to those in the previous chapter, figure 20 shows the coefficient associated with students' disadvantage status in a linear regression model, using their 16-19 attainment as the dependent variable. It demonstrates how the coefficient for disadvantage status changes as an increasing number of other factors are controlled for.

We can see that in terms of raw differences, disadvantaged (but not persistently disadvantaged) students on average achieve more than 2.5 grades less than a non-disadvantaged student (across 3 qualifications). Persistently disadvantaged students (those in receipt of free school meals for over 80 per cent of their time in school) achieve almost 4 grades less.

As we control for other factors (as displayed in the subsequent bars), the combined effect of disadvantage and persistent disadvantage decreases. Most significantly, and consistent with earlier findings, controlling for prior attainment leads to the most sizable reduction in the influence of disadvantage on attainment.

Although the coefficient for disadvantage decreases with the introduction of additional factors, there remains a negative effect associated with disadvantage, of over two fifths of a grade (across three qualifications).

43

Figure 21: Proportion of the 16-19 disadvantage attainment gap that can be explained by differences in the characteristics of disadvantaged and non-disadvantaged students
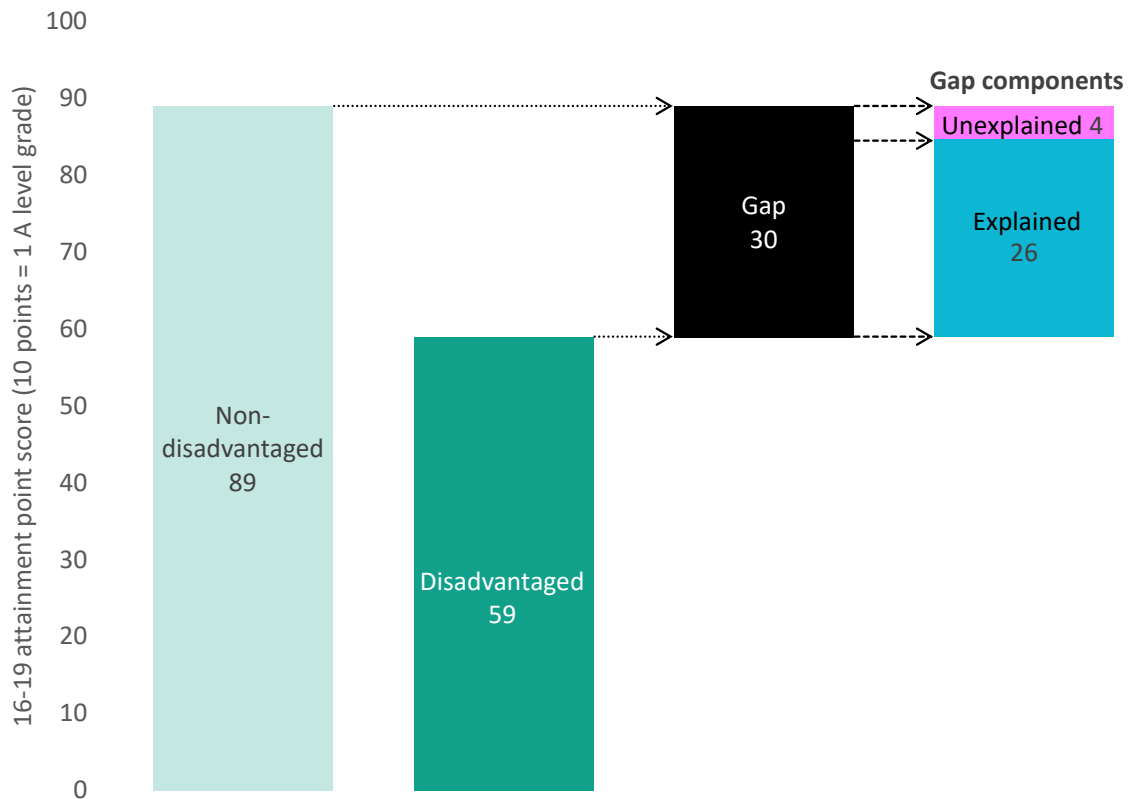
Figure 21 shows that the total disadvantage attainment gap is around 30 points or three A level grades. Eighty six per cent (or just over two and a half A level grades) of this total gap can be explained by differences in the observable characteristics of disadvantaged students and their peers (including the characteristics of the institutions they attend). The remaining 14 per cent (approximately four points, or two fifths of an A level grade) cannot be explained by these characteristics. This indicates that most, but not all, of the 16-19 disadvantage attainment gap is driven by factors other than disadvantage. This unexplained component may be due to other unobserved characteristics or it may be due to disadvantage directly affecting the attainment of 16-19 year olds.

Figure 22: Proportion of the 16-19 disadvantage attainment gap that can be explained by each characteristic
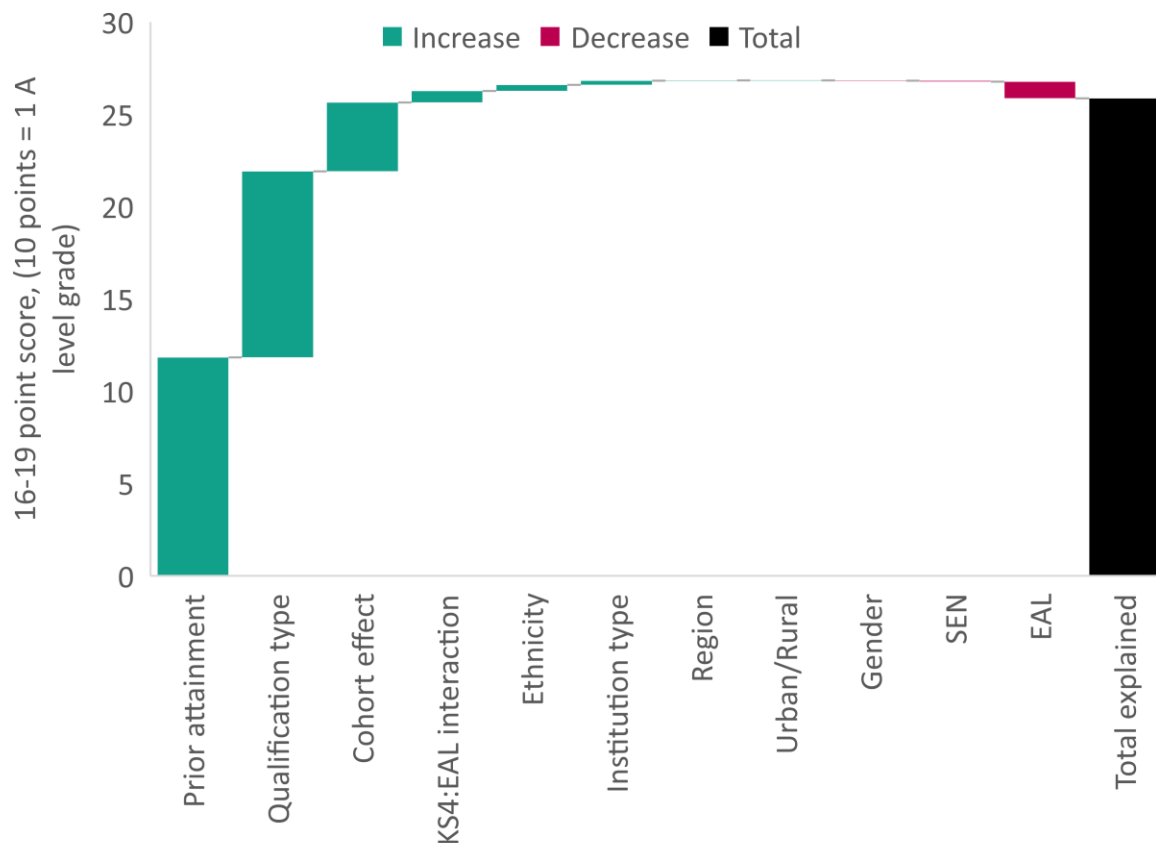
Figure 22 breaks down the explained component of the disadvantage attainment gap (from figure 21) to demonstrate the individual contribution of each characteristic.

We can see from the furthest left bar on the chart that a difference in best three point score of 12, equating to over an A level grade and almost half of the explained gap (26 points), and two fifths of the total gap (30 points), is explained by prior attainment. The ability of students' peers, or the cohort effect, also makes a sizable contribution to the 16-19 disadvantage attainment gap, accounting for two fifths of a grade's difference between the two groups. The type of qualifications entered contributes a grade to the gap. Other demographic characteristics explain an increasingly small amount of the disadvantage attainment gap.

English as an additional language (EAL) makes a negative contribution to the disadvantage attainment gap. This is because, all else being equal, students with EAL have higher 16-19 attainment, but are also more likely to be disadvantaged. This is depicted on the chart by the bar for EAL bringing the cumulative explained variation down, so that the total explained variation is the net impact of all of these factors, some of which (most notably EAL) are negative.

Figure 23: Proportion of the 16-19 disadvantage attainment gap that can be explained by each characteristic factor – using matched Next Steps longitudinal survey data.[13]
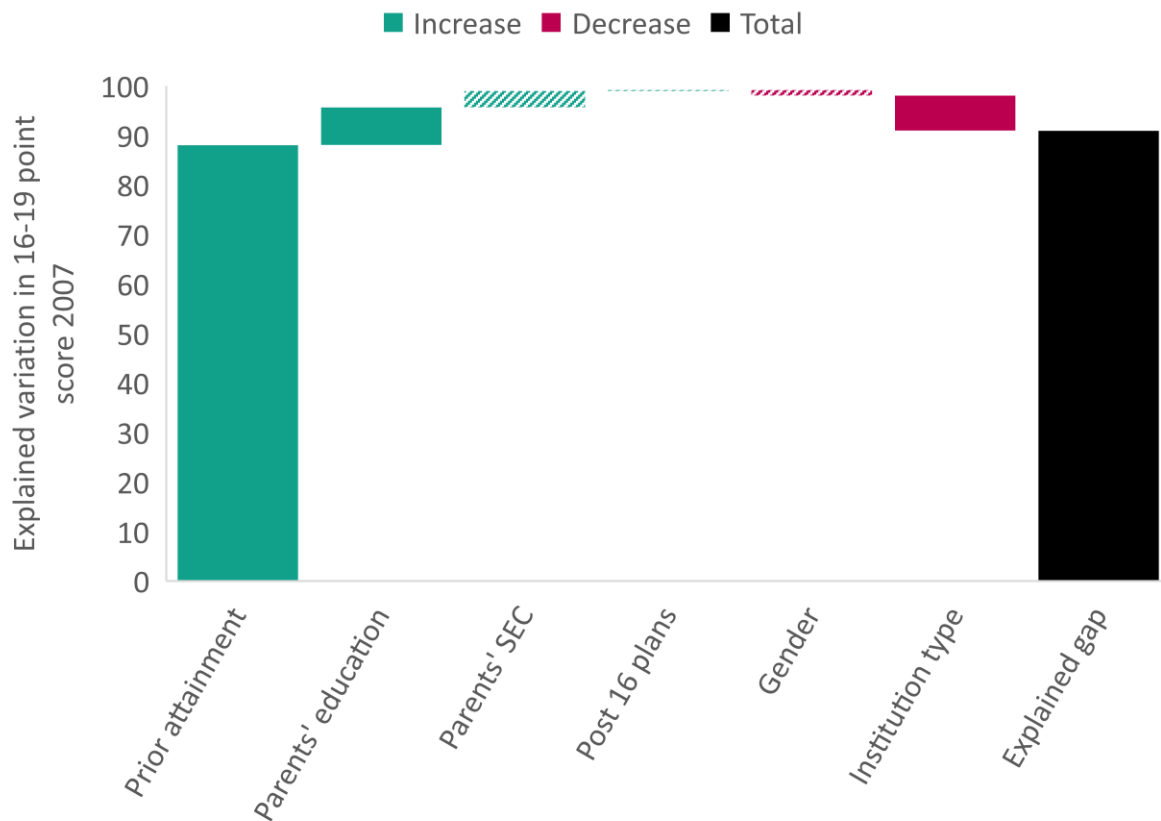
When using variables that capture future plans and parental background, we see that most of the explained gap between the attainment of disadvantaged and non-disadvantaged students can once again be attributed to prior attainment. Parental education levels and socio-economic class and students' post 16 plans do not contribute significantly to the disadvantage attainment gap.

From this finding, we can again conclude that not having these variables available to feed into modelling based on the 2019 data does not present any substantial omission from the analysis.

---

[13] **Source:** University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2020). Next Steps: Linked Education Administrative Datasets (National Pupil Database), England, 2005-2009: Secure Access. [data collection]. 6th Edition. UK Data Service. SN: 7104, http://doi.org/10.5255/UKDA_SN_7104-6

7 - Conclusion

In this report we set out to understand the size of the 16-19 attainment gap, how it has changed in recent years and the factors behind the gap. To do so we have developed a provisional methodology for measures of overall attainment and a measure of how this attainment differs between disadvantaged and non-disadvantaged students.

Combining qualifications as diverse as A levels and level 1 life skills qualifications, along with everything in-between, is challenging. To do so we have had to make a number of assumptions. Whilst it would be tempting to suggest simply that we should instead consider each qualification type separately, that would be a disservice to disadvantaged students. The variety of qualifications available should not serve to obscure the status of these vulnerable young people.

We provide two headline measures; one assuming absolute equivalence between academic and vocational qualifications, another that reflects the greater average returns for academic qualifications. In future, it would be possible to go further to reflect the differential returns of different qualifications and subjects. We acknowledge that there are refinements that could be made to the methodology presented here, and we welcome feedback on our approach.

Using our provisional methodology, we find that the attainment of disadvantaged students is the equivalent of almost three A level grades below that of their better off peers, or over four grades if academic qualifications are given a greater weighting. This gap appears to have fallen only very slightly in recent years.

We investigated a variety of approaches to compare the size of this gap with the size of the gap at the end of secondary school. However, the large number of 16-19 students with no qualification pass grades at level 1 to 3 means the distribution of attainment between the two phases are very different. In addition, many students do not continue into 16-19 education, so direct comparisons risk oversimplifying a complex picture. Our fuller analysis finds that disadvantaged students fall further behind non-disadvantaged students with the same key stage 4 attainment and otherwise similar characteristics. This analysis suggests that economic disadvantage could be contributing almost half a grade to the difference in attainment between these students and their more advantaged peers, beyond the gap already present at the end of the secondary school. For progression to further or higher education, or transitions to the labour market, differences of this size matter.

We find that two fifths of the attainment gap can be explained by student's prior attainment at the end of key stage 4. This has two implications. Firstly, that attempts to close the attainment gap seen at the end of compulsory education should continue to focus on earlier education phases. Secondly, that the plateauing closure of the key stage 4 gap may well feed through into the 16-19 gap once 2020 results are incorporated. That the study and exams of the 2020 cohort were so heavily affected by the impact of the Covid-19 pandemic does not bode well for these students.

Changes to the funding formula for 16-19 education in 2013/14 resulted in a significant increase in the funding targeted towards disadvantaged young people, though this was followed by falls as overall levels of funding fell in real terms. Moreover, students are targeted based on the average level of disadvantage where they live, and not on whether they themselves are economically

disadvantaged. Specifically, the funding formula uses the Index of Multiple Deprivation (IMD) rather than eligibility for free school meals, the measure used to target disadvantaged students up to the age of 16. This may be partly because take-up of free school meals in colleges is low, making the measure less useful in targeting students. However, our analysis shows that free school meal eligibility during secondary school is a marginally better indicator of lower attainment than the IMD measure used in the current funding formula. Using both measures would further improve this targeting.

Given our regression modelling indicates that disadvantaged students fall further behind during the 16-19 phase and that free school meal-based measures have the potential to improve the targeting of funding, we endorse proposals for the introduction of a Student Premium for the 16-19 phase. The introduction of the premium may require a broader review of funding, but crucially overall funding for disadvantaged young people should increase in real terms as a result. Given that average funding rates are low in historic terms, any increase should not be at the expense of other students in this phase. The introduction of the student premium should be accompanied by associated accountability and transparency requirements for providers, to help heighten the focus on disadvantaged students during this phase.

Two other factors make a noticeable contribution to 16-19 attainment and the corresponding gap. Firstly, the ability of students' peers explains over a tenth of the disadvantage gap. Secondly, the type and level of qualification entered explains one third of the gap. These findings are consistent with existing research suggesting that high achieving disadvantaged students are more likely to take qualifications associated with lower prior attainment (academic mismatch, Maragkou 2019). Information, advice and guidance targeted at high attaining disadvantaged students should play a role in addressing this mismatch.

In recent years, the educational inequities suffered by disadvantaged young people have been successfully highlighted by the work of a range of organisations, including EPI. Policymakers should be clear these inequities do not stop once students complete their GCSEs and only start again when some students apply to university.

Annex A: Testing methodology assumptions

In this section we test several key assumptions described in the methodology section of this report.

Number of qualifications included in the 16-19 attainment measure

Our core methodology caps the number of qualifications included in the attainment measure at three. We introduced a cap because we wanted to measure differences in attainment that have a significant real impact on the majority of young people, for example on the likelihood of progression to further or higher education, or into the labour market. There is no evidence that outcomes of the majority of young people would benefit from an excess of qualifications, instead of focussing on a smaller number.

We have decided to use a measure based on three qualifications, as this is the number that represents the most real-world advantage to students in terms of continuing their studies or seeking employment. For example, most entry requirements for higher education have criteria based on three qualifications. This approach also means the trend through time is less sensitive to the decrease in AS level entries discussed previously, as an AS level in addition to a suite of full A levels will not represent substantial real world advantage, for example for University admissions.

However, as many students do take four A levels or equivalent we tested the measure on this basis. When including four qualifications, the disadvantage attainment gap appeared slightly wider. This is because non-disadvantaged students are more likely to take more than three qualifications. It also favoured institutions that encouraged students to enter an additional qualification, such as grammar schools.

Inclusion of level 2 English and maths resits

We have also tested different methodologies for the treatment of English and maths resit qualifications. Specifically, allowing them to count in their entirety, not at all, or only counting progress made since key stage 4.

When counting English and maths resit qualifications in their entirety, the disadvantage attainment gap appears marginally narrower. This is because disadvantaged students are more likely to have taken resit qualifications and therefore benefit from having them included in their best three point score. However, in this scenario, it would be to the relative disadvantage of students who achieved a pass at key stage 4 and therefore were not required to continue their study.

For the same reason, excluding English and maths resit qualifications altogether acts to slightly widen the disadvantage attainment gap.

The differences seen are very minor and do not affect the trend notably through time. However, to treat these qualifications most fairly, the measures presented in this report have been adjusted to only include progress made in English and maths resit qualifications. For example, a student who achieved a grade 3 at key stage 4 and subsequently achieved a grade 5, would have two GCSE grades counting in the best three point score measure, as this represents the improvement they made during their 16-19 study.

It is true that other qualifications taken at age 16 may be retaken during the 16-19 phase. But level 2 English and maths are the only qualifications that are effectively mandated (for all students not achieving a pass during key stage 4), and therefore are considered differently.

The value of level 1/2 v level 3 qualifications

Under method 1 described previously within this report, qualifications of the same level and similar teaching hours are awarded similar points. This does however require an assumption around the weighting of level 1 and 2 qualifications, relative to level 3.

To calculate the initial measure under method 1, we have used AS level points expressed relative to both a GCSE and an A level to create a mapping between level 1/2 and level 3 qualifications. This is described in more detail in Annex C. This mapping results in a grade 9 GCSE (for example), being allocated roughly the same points as a D grade A level or a B grade AS level.

Applying this mapping gives the central line in the chart below, this line has the values displayed in bold and represents the 'method 1 – inputs adjusted' methodology.

Figure 24: Disadvantage attainment gap in A level grades – testing level 1/2 v level 3 weights
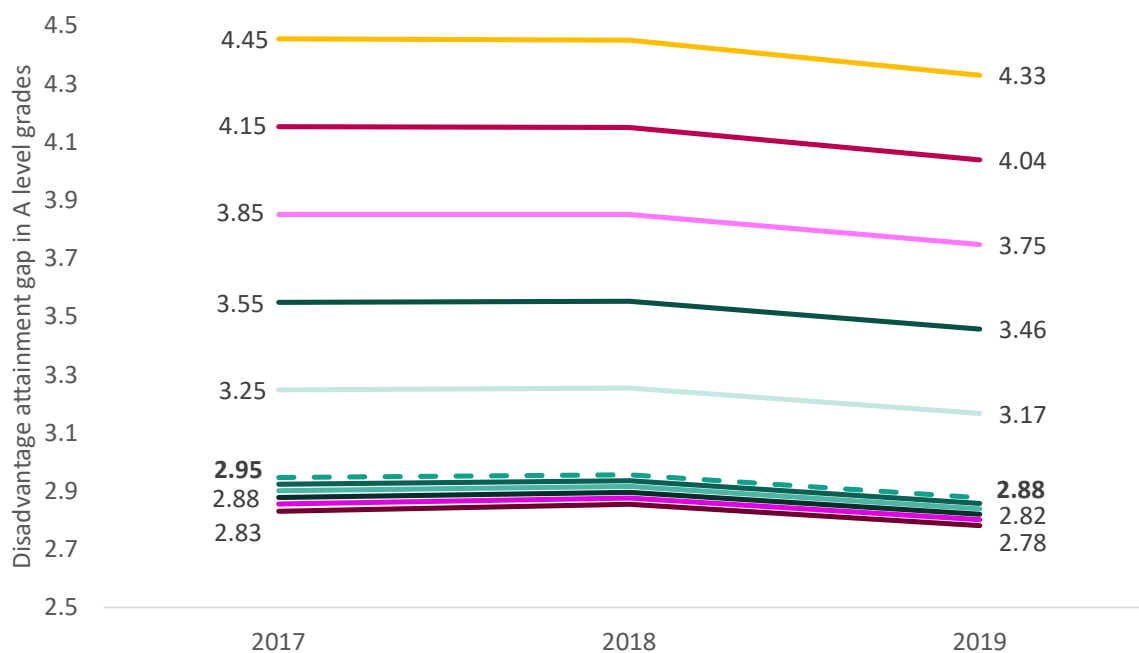
Every line above this, represents the impact of increasing the value of level 3 qualifications in increments of 10 per cent, whilst leaving the points for level 1 and 2 qualifications unchanged. For example, the first line above the one with figures in bold represents the measure if level 3 point scores were multiplied by 1.1. The line above this shows the impact of multiplying level 3 point scores by 1.2, and so on, up to a 50 per cent uplift.

Conversely, every line below this, represents the impact of increasing the value of level 1 and 2 qualifications in increments of 10 per cent up to a 50 per cent uplift, whilst leaving the points for level 3 qualifications unchanged.

We see that increasing the value of level 3 qualifications increases the size of the overall disadvantage attainment gap. This is what we would expect, given that non-disadvantaged students on average, hold more level 3 qualifications.

Conversely, we see that increasing the value of level 1 and 2 qualifications acts to narrow the gap, but only by a very small amount. That is, as the value of level 1 and 2 qualifications are increased, the disadvantage attainment gap moves marginally closer to zero in all years.

Importantly, no matter how we vary this assumption, we see that there is a clear disadvantage attainment gap in 16-19 education and the trend through time is largely unaffected.

The value of level 3 academic vs level 3 vocational and technical qualifications

We also tested the relative value of level 3 academic v level 3 vocational and technical qualifications, by increasing the value of academic qualifications relative to vocational/technical qualifications.

As before, the disadvantage attainment gap with no additional adjustment is displayed on the chart below with the values in bold.

Figure 25: Disadvantage attainment gap in A level grades – testing academic vs vocational weights
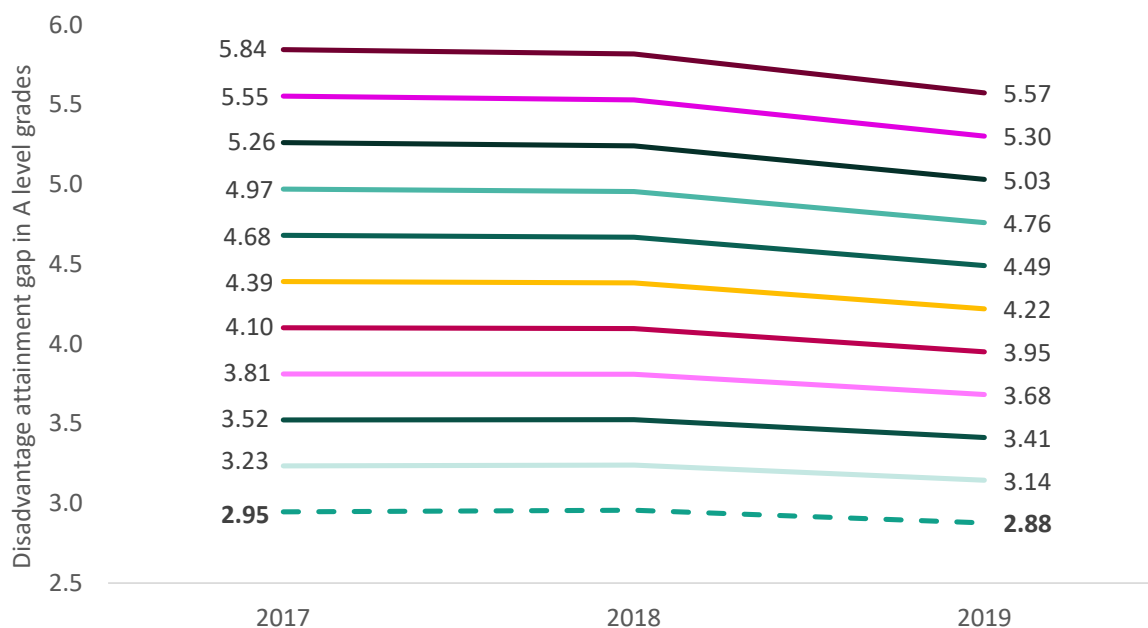
Every line above this, represents the impact of increasing the value of academic qualifications by 10 per cent, up to a total 100 per cent uplift.

In common with the previous chart, we see that increasing the value of academic qualifications widens the overall disadvantage attainment gap.

As with the previous chart, although we see that the gap widens as the value of academic qualifications is inflated, there is no substantial difference in the trend we see through time.

However, although the trend is not sensitive to this assumption, it is known from existing research that students entering academic qualifications have greater labour market returns in later life, meaning there is value in a second measure which adjusts for this. We apply a 46 per cent uplift in method 2, as it represents the additional economic returns to A levels compared to BTECs when taken prior to a first degree (Patrignani, Battiston, and Conlon 2019).

51

Annex B: 16-19 disadvantage attainment gap by Local Authority, 2019

The 16-19 disadvantage attainment gap has been calculated within each local authority as the equivalent number of A level grades that disadvantaged students within each local authority are behind non-disadvantaged students nationally. Figures for the City of London and Rutland have been suppressed due to small student numbers.

Local authority	Disadvantage attainment gap, A level grades	Number of disadvantaged students
Knowsley	5.4	141
North Somerset	4.8	459
Stockton-on-Tees	4.7	291
Torbay	4.4	357
Swindon	4.4	466
Derby	4.3	880
Barnsley	4.3	652
Hartlepool	4.2	238
South Gloucestershire	4.1	556
West Sussex	4.1	1,161
Wiltshire	4.0	554
Southend-on-Sea	4.0	785
Worcestershire	4.0	636
Isle of Wight	4.0	240
Rotherham	4.0	819
Newcastle upon Tyne	4.0	1,845
Nottinghamshire	3.9	1,003
Southampton	3.9	493
Norfolk	3.9	1,497
Warwickshire	3.9	1,065
Medway	3.9	712
North East Lincolnshire	3.9	633
Stoke-on-Trent	3.9	537
Telford and Wrekin	3.8	399
Portsmouth	3.8	356
Nottingham	3.8	1,282
Somerset	3.7	822
Walsall	3.7	1,235
East Sussex	3.7	887
County Durham	3.7	1,139
Central Bedfordshire	3.7	308
York	3.7	547
Suffolk	3.7	1,279

Local authority	Disadvantage attainment gap, A level grades	Number of disadvantaged students
Oxfordshire	3.7	1,094
Plymouth	3.7	602
Middlesbrough	3.7	830
Blackpool	3.7	663
Bournemouth, Christchurch and Poole	3.7	590
South Tyneside	3.6	305
Kent	3.6	2,690
Sheffield	3.6	1,267
Sefton	3.6	712
North Lincolnshire	3.6	555
Salford	3.6	629
Bedford	3.6	645
Cheshire East	3.6	741
Dorset	3.5	440
Wakefield	3.5	779
Peterborough	3.5	666
Staffordshire	3.5	1,428
Thurrock	3.5	77
Lincolnshire	3.5	1,105
Kingston upon Hull, City of	3.5	863
Cumbria	3.4	587
Leeds	3.4	2,040
Bromley	3.4	1,021
Northamptonshire	3.4	1,056
Northumberland	3.4	377
East Riding of Yorkshire	3.4	404
Bristol, City of	3.4	772
Gateshead	3.4	468
Cambridgeshire	3.4	729
Kingston upon Thames	3.4	808
Coventry	3.3	832
Barking and Dagenham	3.3	927
Blackburn with Darwen	3.3	462
Milton Keynes	3.3	691
Halton	3.3	682
Dudley	3.3	1,203
Hampshire	3.2	2,068
Hammersmith and Fulham	3.2	685
Rochdale	3.2	738
Richmond upon Thames	3.1	318

Local authority	Disadvantage attainment gap, A level grades	Number of disadvantaged students
St. Helens	3.1	645
Essex	3.1	2,426
Sunderland	3.1	719
Bradford	3.1	1,364
Trafford	3.1	424
Havering	3.1	728
Leicestershire	3.0	723
West Berkshire	3.0	169
Surrey	3.0	1,402
Bath and North East Somerset	3.0	280
Hertfordshire	3.0	2,231
Herefordshire, County of	3.0	293
Bracknell Forest	3.0	127
Bolton	3.0	775
Windsor and Maidenhead	2.9	198
Cheshire West and Chester	2.9	276
Devon	2.9	932
Doncaster	2.9	424
Derbyshire	2.9	653
Solihull	2.9	1,225
Buckinghamshire	2.9	504
Wigan	2.9	707
Cornwall	2.8	1,021
Shropshire	2.8	362
Warrington	2.8	396
Calderdale	2.8	463
Stockport	2.8	415
Tameside	2.8	668
Gloucestershire	2.7	719
Hillingdon	2.7	1,186
Lancashire	2.7	1,943
Wirral	2.7	948
Darlington	2.7	389
Camden	2.6	2,206
Sandwell	2.6	1,236
North Tyneside	2.5	139
Kirklees	2.5	938
Leicester	2.5	1,043
Wolverhampton	2.5	693
Greenwich	2.4	680

Local authority	Disadvantage attainment gap, A level grades	Number of disadvantaged students
Slough	2.3	600
Bury	2.3	709
Liverpool	2.3	1,398
Haringey	2.2	667
North Yorkshire	2.2	529
Birmingham	2.1	4,211
Oldham	2.0	798
Waltham Forest	2.0	1,462
Manchester	2.0	2,060
Lambeth	2.0	609
Croydon	1.8	945
Brighton and Hove	1.8	399
Lewisham	1.5	717
Luton	1.4	607
Tower Hamlets	1.2	1,625
Redcar and Cleveland	1.2	230
Hounslow	1.2	745
Wandsworth	1.2	836
Westminster	1.1	1,192
Enfield	1.0	711
Reading	1.0	96
Kensington and Chelsea	1.0	494
Barnet	1.0	1,159
Brent	0.4	449
Hackney	0.2	713
Wokingham	0.1	82
Harrow	0.1	617
Bexley	0.0	238
Islington	-0.1	293
Newham	-0.2	1,622
Merton	-0.2	160
Sutton	-0.2	297
Ealing	-0.5	568
Redbridge	-0.5	725
Southwark	-1.2	433

Annex C: Technical documentation

Underlying datasets and inclusion criteria

The datasets used by EPI to produce the measures presented in this report are provided via the Department for Education (DfE) and accessed via the Office for National Statistics' Secure Research Service.

The student level National Pupil Database (NPD) is used to identify all students at the end of 16-19 study who were affiliated with a main-stream school or college. This is a composite database including those who were in a sixth form as recorded in the school census, and those who were enrolled or took qualifications at FE colleges or other organisations which complete an Individualised Learner Record (ILR) return. We further remove apprentices from our analysis due to the difficulty in allocating them a 'qualification grade'.

The exam level NPD has been used to identify the level 1-3 qualifications that these students entered in the two or three year period since finishing key stage 4. From 2017 onwards the exam level NPD includes all regulated qualifications up to level 3 (as listed in the Ofqual qualification register) entered by students of the relevant age. This includes qualifications under the Qualifications Credit Framework, which sits within the regulated qualification register.

We further remove any English or maths entries by students that were obliged to continue study of these subjects because of the English and maths condition of funding policy. However, where students have been obliged to continue study of these subjects and have made positive progress since the end of key stage 4, we create an exam record with points equal to the amount of progress they have made, rather than the overall grade.

Students in a maintained institution who are at the end of their study but have no level 1-3 qualifications will remain in the measure, with a point score allocation of zero.

All level 1-3 qualifications are included, regardless of grade scheme. For example, pass/fail qualifications or those with any other grade scheme are in scope and will have points allocated as set out below.

Point score allocations

For the purpose of this project, the starting assumption when allocating point scores is to assume equivalence based on qualification level, guided learning hours and grade scheme.

For example, all level 2 Pass/Merit/Distinction qualifications with the same number of guided learning hours will be awarded the same points.

The same would be true for a level 2 qualification with the same guided learning hours and a C/B/A grade scheme – the important thing is the number of distinct pass grades available rather than what name these grades are given.

Level 3 qualifications

For a large number of the qualifications in scope, the points will be the same ones that are used in performance tables measures, created by the Department for Education and described in their

published guidance.[14] This is extended to all level 3 qualifications in the exam level dataset, regardless of whether the qualification is eligible to count in the performance tables.

Level 2 qualifications approved for 16-19 study

For lower-level qualifications, those that count in the 16-19 performance tables will again have the same points used by the Department for Education. These points are extended to all level 2 qualifications approved for 16-19 study. The potential points available from a level 2 qualification will always be less than the potential points available from a level 3 qualification of the same size.

Other level 1 and 2 qualifications

For other level 1 and 2 qualifications, we have set points in a similar way based on level, guided learning hours and grade scheme. However, as an interim step points are expressed relative to a 9-1 GCSE as set out in Annex B of the Department for Education's[15] secondary accountability guidance, before a secondary mapping is applied to rescale them relative to an A level.

This approach is applied to all level 1 and 2 qualifications which do not already have points attributed to them on the correct scale, regardless of whether these qualifications are eligible to count towards the key stage 4 performance tables.

This approach to setting point scores is intended for qualifications of equivalent size (in terms of guided learning hours) to a GCSE. Where level 1 and 2 qualifications have been sat with less guided learning hours than a GCSE, points are set on the same basis but then multiplied through by their size equivalence. For example, a qualification with one quarter of the guided learning hours of a GCSE would have points set as described above, then multiplied by 0.25.

Once all remaining level 1 and 2 qualifications have had points set on this basis, we use the fact that an AS level has points expressed under both systems, that is, relative to both a GCSE and an A level, to create a mapping.

---

[14] Performance points – a practical guide to key stage 4 and 16 to 18 performance points
[15] Secondary accountability measures guidance February 2020

Points awarded to each AS level grade, in key stage 4 reporting relative to a GCSE and key stage 5 reporting relative to an A level
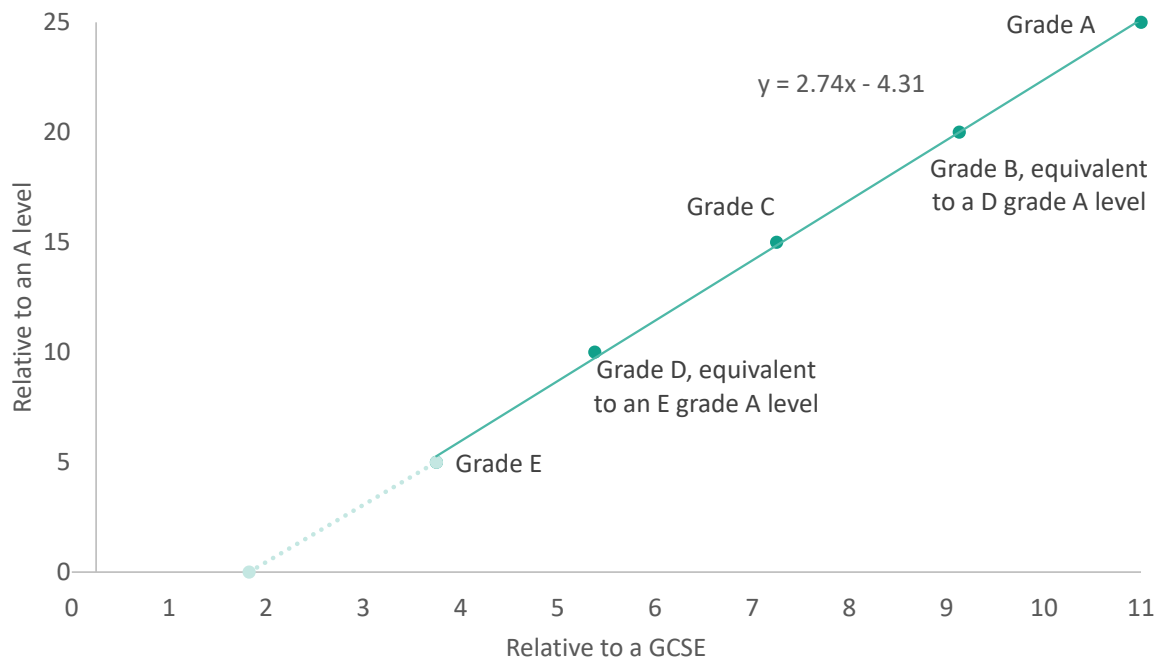
Plotting a line of best fit on the chart above yields the equation y=2.74x-4.31. We make an assumption that this relationship can be extrapolated to the full range of level 1 and 2 point scores available, to map the remaining qualifications onto a point scale relative to a qualification with similar teaching hours to an A level.

As a final step, we shift the value of all qualifications (not just level 1 and 2) up by 4.31 (meaning the intercept on the chart above becomes zero). This is for presentational purposes only, it serves to maintain the relative distance between qualifications and grades, but ensures all qualifications attract positive points. On this basis, a difference of 10 points, can be seen as equivalent to 1 A level grade.

Worked example

Consider a level 2 qualification with a pass/merit/distinction grade scheme and similar guided learning hours to a GCSE. The important consideration is not what the grades are called, but how many distinct pass grades there are.

In this case, as there are three distinct pass grades, we consult the table on page 38 of the Department for Education's guidance which gives the point score table below.

Level 2 grade structure	Example grade	2016 points	2017 to 2019 points
4 grade scheme	A*	8.00	8.50
	A	7.00	7.00
	B	6.00	5.50
	C	5.00	4.00
3 grade scheme	Distinction	7.50	7.75
	Merit	6.50	6.25
	Pass	5.00	4.00
Pass only	Pass	6.00	5.50

In this case, the points expressed relative to a 9-1 GCSE (2017-2019 points) for the Pass/Merit/Distinction grades are 4.00, 6.25 and 7.75 respectively.

We then apply the mapping based on the line of best fit discussed above, that is y=2.74x-4.31, to map these points onto a scale relative to an A level.

This transformation gives 6.65, 12.815 and 16.925 for the three respective grades, which we then shift up by 4.31 to obtain 10.96, 17.125 and 21.235.

For reference, a D grade A level under this system would attract 24.31 points. This means that in this example, a distinction grade in a level 2 pass/merit/distinction qualification would be awarded points just below the equivalent of a D grade A level. This compares to a grade 9 GCSE which would be allocated 24.66 points, also roughly equivalent to a D grade A level.

Note that if this qualification only had half the guided learning hours of a GCSE, we would have multiplied the points relative to a GCSE by 0.5 prior to applying the mapping derived from an AS level.

Student's best three results are then calculated based on the size of their qualifications rather than the distinct number. Where qualifications of size equal to a GCSE are considered 0.25 the size of an A level. As an AS level is of size 0.5 relative to an A level, students could therefore have two A levels, an AS level and two GCSEs counting as their best three qualifications if these were all taken during 16-19 study.

Qualification discounting has been applied such that if students enter the same qualification (or a different qualification with substantial overlap in content) multiple times, only one result is eligible to count in their best three point score.
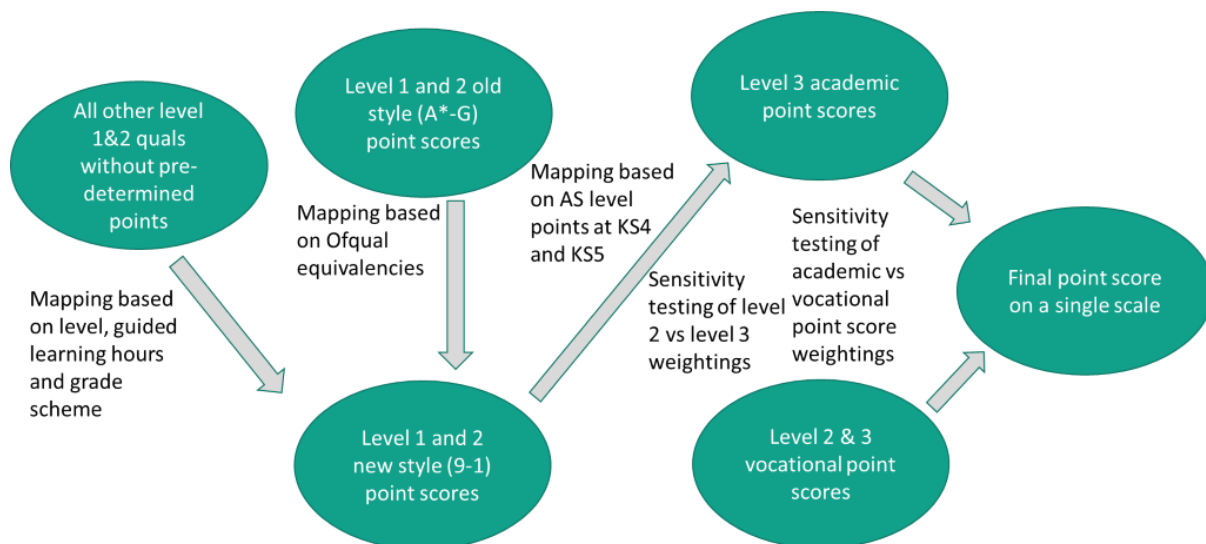
The flow diagram above depicts the full process for all level 1, 2 and 3 qualifications. It also indicates the points where sensitivity testing has been applied to test the robustness of assumptions.

Once a total point score has been calculated for each student based on their best three results, the average of this score is taken separately for non-disadvantaged and disadvantaged students. This difference between the mean point score for each group is then divided by 10 (the difference in points between each A level grade) to calculate the overall disadvantage attainment gap, expressed as an average number of A level grades.

Annex D: Full results from 16-19 attainment regression models

The models presented in the tables below are as follows, where model 1-6 build iteratively on each other, all models have 16-19 best three point score, method 1 as the dependent variable:

- Model 1: Just disadvantage and persistent disadvantage
- Model 2: + student demographics
- Model 3: + prior attainment
- Model 4: + geographic factors
- Model 5: + main qualification type
- Model 6: + institution type
- Model 7: full model without prior attainment
- Model 8: model 5 without IMD
- Model 9: model 6 without IMD
- Model 10: model 6 without disadvantage, persistent disadvantage and IMD and with binary prior attainment
- Model 11: model 6 without disadvantage and persistent disadvantage and with bottom 27% IMD
- Model 12: model 6 without IMD and persistent disadvantage and with binary prior attainment
- IMD and binary prior attainment
- Model 13: model 6 without persistent disadvantage and with bottom 27% IMD and binary prior attainment

For the purpose of these models, the groupings below are implemented:

Student demographics:

- Gender
- Ethnicity
- English as a second language
- Special educational needs

Prior attainment:

- Average key stage 4 English and maths results
- Average key stage 4 English and maths results squared
- Difference in key stage 4 English and maths results
- Interaction term between average key stage 4 English and maths and English as a second language status

Geographics variables:

- Urban or rural
- Government Office Region

- Index of multiple deprivation
- Cohort effect (prior attainment of cohort at the same institution, also removed in model 6)

The reference category for institution type is local authority-maintained schools
The reference category for main qualification pathway is academic
The reference category for region is London
The reference category for ethnicity is White British

Dependent variable is best 3 point score, method 1 throughout	Coefficients and standard errors					
Independent variables	Model 1	SE	Model 2	SE	Model 3	SE
(Intercept)	85.16 *	0.08	82.61 *	0.12	14.11 *	0.46
Disadvantage	-24.49 *	0.19	-25.85 *	0.19	-8.27 *	0.15
Persistent disadvantage	-13.45 *	0.30	-8.77 *	0.30	-1.08 *	0.23
Gender: Female			12.30 *	0.14	5.22 *	0.11
Special Educational Needs (SEN)			-30.55 *	0.22	-1.29 *	0.18
English as an Additional Language (EAL)			-5.02 *	0.29	13.81 *	0.54
Ethnicity: Any other Asian background			14.59 *	0.57	-0.48	0.42
Ethnicity: Any other black background			1.11	0.85	2.72 *	0.64
Ethnicity: Any other ethnic group			12.49 *	0.59	3.63 *	0.44
Ethnicity: Any other mixed background			8.92 *	0.55	1.23 *	0.41
Ethnicity: Any other white background			3.93 *	0.39	1.57 *	0.29
Ethnicity: Bangladeshi			17.28 *	0.57	1.88 *	0.43
Ethnicity: Black - African			11.01 *	0.41	2.41 *	0.31
Ethnicity: Black Caribbean			-6.62 *	0.59	-0.57	0.44
Ethnicity: Chinese			35.82 *	1.08	8.70 *	0.80
Ethnicity: Gypsy/Roma			-38.46 *	1.84	-10.73 *	1.51
Ethnicity: Indian			19.91 *	0.46	3.70 *	0.34
Ethnicity: Information not yet obtained			1.13	0.99	1.53 *	0.75
Ethnicity: Pakistani			5.48 *	0.40	1.51 *	0.30
Ethnicity: Refused			5.92 *	1.00	0.64	0.74
Ethnicity: Traveller of Irish Heritage			-24.61 *	5.86	-8.49	4.66
Ethnicity: White - Irish			12.49 *	1.24	1.93 *	0.92
Ethnicity: White and Asian			13.65 *	0.71	1.56 *	0.53
Ethnicity: White and Black African			6.64 *	0.96	0.16	0.71
Ethnicity: White and Black Caribbean			-5.69 *	0.62	-1.99 *	0.46
Average of key stage 4 English and maths					-5.28 *	0.17
Average of key stage 4 English and maths squared					2.90 *	0.02
Difference between key stage 4 English and maths results					-1.87 *	0.05
EAL:Average of key stage 4 English and maths interaction term					-2.13 *	0.09
R squared	0.059		0.138		0.523	

* indicates significance at the 95% level

Dependent variable is best 3 point score, method 1 throughout	Coefficients and standard errors					
Independent variables	Model 4	SE	Model 5	SE	Model 6	SE
(Intercept)	-26.99 *	0.63	53.99 *	0.67	58.19 *	0.81
Disadvantage	-5.98 *	0.15	-3.67 *	0.13	-3.69 *	0.13
Persistent disadvantage	-0.44 *	0.22	0.43 *	0.20	0.44 *	0.20
Gender: Female	5.17 *	0.11	3.67 *	0.10	3.66 *	0.10
Special Educational Needs (SEN)	-0.79 *	0.18	0.62 *	0.16	0.63 *	0.16
English as an Additional Language (EAL)	15.23 *	0.53	9.54 *	0.48	9.20 *	0.48
Ethnicity: Any other Asian background	-2.11 *	0.42	-2.95 *	0.38	-2.91 *	0.38
Ethnicity: Any other black background	2.78 *	0.63	-0.76	0.57	-0.91	0.57
Ethnicity: Any other ethnic group	2.73 *	0.43	0.89 *	0.39	0.92 *	0.39
Ethnicity: Any other mixed background	0.89 *	0.40	0.02	0.37	-0.01	0.37
Ethnicity: Any other white background	1.96 *	0.29	1.41 *	0.26	1.44 *	0.26
Ethnicity: Bangladeshi	1.46 *	0.42	-0.53	0.38	-0.59	0.38
Ethnicity: Black - African	1.90 *	0.31	-1.49 *	0.28	-1.60 *	0.28
Ethnicity: Black Caribbean	1.35 *	0.44	-1.31 *	0.40	-1.53 *	0.40
Ethnicity: Chinese	6.15 *	0.79	5.99 *	0.71	6.07 *	0.71
Ethnicity: Gypsy/Roma	-9.01 *	1.47	-6.43 *	1.34	-6.33 *	1.34
Ethnicity: Indian	2.46 *	0.33	1.31 *	0.30	1.28 *	0.30
Ethnicity: Information not yet obtained	0.96	0.73	1.32 *	0.66	1.19	0.66
Ethnicity: Pakistani	0.86 *	0.30	-0.58 *	0.27	-0.66 *	0.27
Ethnicity: Refused	0.01	0.73	-0.52	0.66	-0.49	0.66
Ethnicity: Traveller of Irish Heritage	-8.82	4.55	-3.03	4.13	-3.05	4.13
Ethnicity: White - Irish	0.84	0.90	0.77	0.82	0.80	0.82
Ethnicity: White and Asian	0.39	0.51	0.23	0.47	0.23	0.47
Ethnicity: White and Black African	-0.03	0.70	-1.45 *	0.63	-1.50 *	0.63
Ethnicity: White and Black Caribbean	-0.73	0.46	-1.22 *	0.41	-1.28 *	0.41
Average of key stage 4 English and maths	-4.41 *	0.17	-20.03 *	0.16	-20.20 *	0.16
Average of key stage 4 English and maths squared	2.37 *	0.02	3.30 *	0.02	3.32 *	0.02
Difference between key stage 4 English and maths results	-1.55 *	0.05	-1.06 *	0.04	-1.06 *	0.04
EAL:Average of key stage 4 English and maths interaction term	-2.24 *	0.09	-1.43 *	0.08	-1.38 *	0.08
School or college in an urban area	-0.49 *	0.21	0.32	0.19	0.18	0.19
Region: East Midlands	-1.14 *	0.24	-0.05	0.22	0.12	0.22
Region: East of England	0.58 *	0.21	1.03 *	0.19	1.18 *	0.20
Region: North East	4.30 *	0.28	3.27 *	0.26	3.55 *	0.26
Region: North West	6.00 *	0.21	5.33 *	0.19	5.28 *	0.19
Region: South East	-0.34	0.20	0.67 *	0.18	0.71 *	0.18
Region: South West	0.84 *	0.23	1.88 *	0.21	2.17 *	0.21
Region: West Midlands	1.24 *	0.21	1.02 *	0.19	1.10 *	0.20
Region Yorkshire and the Humber	2.97 *	0.23	2.53 *	0.21	2.51 *	0.21
Index of Multiple Deprivation	-0.11 *	0.00	-0.08 *	0.00	-0.08 *	0.00

Dependent variable is best 3 point score, method 1 throughout	Coefficients and standard errors					
Independent variables	Model 4	SE	Model 5	SE	Model 6	SE
Average KS4 English and maths of cohort	9.91 *	0.07	7.01 *	0.07	6.38 *	0.10
Main qualification type: GCSE			-48.79 *	0.36	-48.75 *	0.36
Main qualification type: Other level 1&2			-45.84 *	0.20	-45.58 *	0.20
Main qualification type: Non-academic level 3			-0.07	0.15	0.07	0.15
Institution type: Academy					-0.12	0.17
Institution type: Colleges					-1.41 *	0.22
Institution type: Free schools					-0.11	0.37
Institution type: Other					-4.01	3.08
Institution type: Sixth form college					1.95 *	0.21
Institution type: Special schools					6.76 *	1.69
R squared	0.546		0.625		0.626	

* indicates significance at the 95% level

Dependent variable is best 3 point score, method 1 throughout	Coefficients and standard errors					
Independent variables	Model 7	SE	Model 8	SE	Model 9	SE
(Intercept)	111.73 *	0.31	51.80 *	0.66	55.08 *	0.79
Disadvantage	-6.63 *	0.14	-4.21 *	0.13	-4.24 *	0.13
Persistent disadvantage	0.32	0.22	0.10	0.20	0.11	0.20
Gender: Female	4.73 *	0.10	3.61 *	0.10	3.59 *	0.10
Special Educational Needs (SEN)	-3.77 *	0.17	0.65 *	0.16	0.66 *	0.16
English as an Additional Language (EAL)	-1.26 *	0.21	9.63 *	0.48	9.35 *	0.48
Ethnicity: Any other Asian background	1.24 *	0.42	-3.32 *	0.38	-3.29 *	0.38
Ethnicity: Any other black background	-3.70 *	0.62	-1.42 *	0.57	-1.57 *	0.57
Ethnicity: Any other ethnic group	1.89 *	0.43	0.43	0.39	0.44	0.39
Ethnicity: Any other mixed background	1.13 *	0.40	-0.27	0.37	-0.30	0.37
Ethnicity: Any other white background	1.94 *	0.29	0.99 *	0.26	1.00 *	0.26
Ethnicity: Bangladeshi	1.89 *	0.42	-1.27 *	0.38	-1.34 *	0.38
Ethnicity: Black - African	-2.09 *	0.31	-2.25 *	0.28	-2.37 *	0.28
Ethnicity: Black Caribbean	-5.45 *	0.44	-1.95 *	0.40	-2.18 *	0.40
Ethnicity: Chinese	15.62 *	0.79	5.80 *	0.71	5.86 *	0.71
Ethnicity: Gypsy/Roma	-4.46 *	1.34	-6.93 *	1.34	-6.84 *	1.34
Ethnicity: Indian	5.69 *	0.33	1.02 *	0.30	0.98 *	0.30
Ethnicity: Information not yet obtained	1.31	0.72	1.02	0.66	0.89	0.66
Ethnicity: Pakistani	-0.16	0.30	-1.32 *	0.27	-1.41 *	0.27
Ethnicity: Refused	0.25	0.72	-0.69	0.66	-0.67	0.66
Ethnicity: Traveller of Irish Heritage	-3.84	4.24	-3.16	4.14	-3.18	4.13
Ethnicity: White - Irish	2.64 *	0.90	0.62	0.82	0.64	0.81
Ethnicity: White and Asian	3.68 *	0.51	0.15	0.47	0.14	0.47

Dependent variable is best 3 point score, method 1 throughout			Coefficients and standard errors				
Independent variables	Model 7		SE	Model 8		SE	Model 9		SE

Independent variables	Model 7	SE	Model 8	SE	Model 9	SE
Ethnicity: White and Black African	-1.14	0.70	-1.87 *	0.63	-1.93 *	0.63
Ethnicity: White and Black Caribbean	-3.11 *	0.45	-1.56 *	0.41	-1.63 *	0.41
Average of key stage 4 English and maths			-19.97 *	0.16	-20.11 *	0.16
Average of key stage 4 English and maths squared			3.30 *	0.02	3.32 *	0.02
Difference between key stage 4 English and maths results			-1.08 *	0.04	-1.07 *	0.04
EAL:Average of key stage 4 English and maths interaction term			-1.49 *	0.08	-1.45 *	0.08
School or college in an urban area	1.53 *	0.21	0.09	0.19	-0.08	0.19
Region: East Midlands	-2.09 *	0.24	-0.23	0.22	-0.09	0.22
Region: East of England	-0.80 *	0.21	1.08 *	0.19	1.21 *	0.20
Region: North East	3.32 *	0.28	2.58 *	0.26	2.79 *	0.26
Region: North West	4.92 *	0.21	4.58 *	0.19	4.48 *	0.19
Region: South East	-0.38	0.20	0.79 *	0.18	0.81 *	0.18
Region: South West	1.27 *	0.23	1.77 *	0.21	2.02 *	0.21
Region: West Midlands	-0.15	0.21	0.53 *	0.19	0.57 *	0.19
Region Yorkshire and the Humber	1.24 *	0.23	1.92 *	0.21	1.88 *	0.21
Index of Multiple Deprivation	-0.19 *	0.00				
Average KS4 English and maths of cohort			7.18 *	0.07	6.68 *	0.10
Main qualification type: GCSE	-80.89 *	0.36	-48.85 *	0.36	-48.85 *	0.36
Main qualification type: Other level 1&2	-78.84 *	0.18	-45.95 *	0.20	-45.75 *	0.20
Main qualification type: Non-academic level 3	-24.67 *	0.14	-0.08	0.15	0.02	0.15
Institution type: Academy	2.05 *	0.19			-0.09	0.17
Institution type: Colleges	-11.29 *	0.21			-1.06 *	0.22
Institution type: Free schools	3.55 *	0.41			-0.18	0.37
Institution type: Other	-14.47 *	2.98			-3.56	3.08
Institution type: Sixth form college	2.20 *	0.23			1.99 *	0.21
Institution type: Special schools	-22.06 *	0.66			7.71 *	1.69
R squared	0.548		0.625		0.625	

* indicates significance at the 95% level

Dependent variable is best 3 point score, method 1 throughout			Coefficients and standard errors				
Independent variables	Model 10	SE	Model 11	SE	Model 12	SE
---	---	---	---	---	---	---
(Intercept)	98.28 *	0.39	100.27 *	0.39	98.32 *	0.39
Disadvantage			-7.84 *	0.13		
IMD rank for bottom 27%					-30.80 *	0.74
Gender: Female	4.49 *	0.10	4.73 *	0.10	4.60 *	0.10
Special Educational Needs (SEN)	-2.70 *	0.17	-2.30 *	0.17	-2.62 *	0.17
English as an Additional Language (EAL)	-2.09 *	0.21	-1.76 *	0.21	-1.52 *	0.22
Ethnicity: Any other Asian background	0.39	0.42	0.49	0.41	0.78	0.42

Dependent variable is best 3 point score, method 1 throughout	Coefficients and standard errors					
Independent variables	Model 10	SE	Model 11	SE	Model 12	SE
Ethnicity: Any other black background	-6.43 *	0.62	-4.83 *	0.62	-4.95 *	0.62
Ethnicity: Any other ethnic group	-0.25	0.43	1.07 *	0.43	0.74	0.43
Ethnicity: Any other mixed background	-0.25	0.40	0.53	0.40	0.33	0.40
Ethnicity: Any other white background	1.49 *	0.29	1.09 *	0.29	2.12 *	0.29
Ethnicity: Bangladeshi	-1.62 *	0.42	0.08	0.42	0.27	0.42
Ethnicity: Black - African	-5.59 *	0.31	-3.69 *	0.31	-3.90 *	0.31
Ethnicity: Black Caribbean	-7.96 *	0.44	-6.74 *	0.44	-6.51 *	0.44
Ethnicity: Chinese	15.39 *	0.79	15.03 *	0.78	15.64 *	0.79
Ethnicity: Gypsy/Roma	-5.02 *	1.34	-3.64 *	1.34	-3.69 *	1.34
Ethnicity: Indian	5.72 *	0.33	5.17 *	0.33	5.79 *	0.33
Ethnicity: Information not yet obtained	0.57	0.72	0.88	0.72	1.07	0.72
Ethnicity: Pakistani	-2.51 *	0.30	-1.76 *	0.29	-1.15 *	0.30
Ethnicity: Refused	-0.48	0.73	0.01	0.72	-0.16	0.73
Ethnicity: Traveller of Irish Heritage	-5.52	4.26	-3.25	4.24	-4.87	4.25
Ethnicity: White - Irish	2.52 *	0.90	2.43 *	0.90	2.67 *	0.90
Ethnicity: White and Asian	3.23 *	0.52	3.62 *	0.51	3.41 *	0.52
Ethnicity: White and Black African	-3.29 *	0.70	-1.92 *	0.70	-2.47 *	0.70
Ethnicity: White and Black Caribbean	-5.09 *	0.45	-3.83 *	0.45	-4.29 *	0.45
Achieved grade 4 in English prior attainment	2.76 *	0.23	2.33 *	0.23	2.49 *	0.23
Achieved grade 4 in maths prior attainment	7.62 *	0.19	7.04 *	0.19	7.39 *	0.19
School or college in an urban area	0.57 *	0.21	0.88 *	0.21	1.06 *	0.21
Region: East Midlands	-1.82 *	0.24	-2.51 *	0.24	-1.38 *	0.24
Region: East of England	0.20	0.21	-0.62 *	0.21	0.09	0.21
Region: North East	1.84 *	0.28	1.70 *	0.28	3.00 *	0.28
Region: North West	3.40 *	0.21	3.14 *	0.21	4.76 *	0.21
Region: South East	0.88 *	0.20	-0.03	0.20	0.69 *	0.20
Region: South West	1.81 *	0.23	1.01 *	0.23	1.91 *	0.23
Region: West Midlands	-1.02 *	0.21	-1.33 *	0.21	0.00	0.22
Region Yorkshire and the Humber	0.25	0.23	-0.18	0.23	1.42 *	0.23
Main qualification type: GCSE	-79.27 *	0.37	-77.99 *	0.37	-78.58 *	0.37
Main qualification type: Other level 1&2	-78.18 *	0.19	-76.56 *	0.19	-77.37 *	0.19
Main qualification type: Non-academic level 3	-25.58 *	0.14	-24.98 *	0.14	-25.16 *	0.14
Institution type: Academy	2.29 *	0.19	2.24 *	0.19	2.26 *	0.19
Institution type: Colleges	-11.12 *	0.21	-10.82 *	0.21	-11.05 *	0.21
Institution type: Free schools	2.87 *	0.41	3.18 *	0.41	3.07 *	0.41
Institution type: Other	-12.55 *	3.00	-12.95 *	2.98	-12.57 *	2.99
Institution type: Sixth form college	2.15 *	0.23	2.20 *	0.23	2.15 *	0.23
Institution type: Special schools	-16.88 *	0.68	-16.86 *	0.68	-17.25 *	0.68
R squared	0.544		0.548		0.546	

* indicates significance at the 95% level

Dependent variable is best 3 point score, method 1 throughout	Coefficients and standard errors	
Independent variables	Model 13	SE
(Intercept)	100.11 *	0.39
Disadvantage	-7.02 *	0.13
IMD rank for bottom 27%	-21.45 *	0.76
Gender: Female	4.79 *	0.10
Special Educational Needs (SEN)	-2.28 *	0.17
English as an Additional Language (EAL)	-1.40 *	0.21
Ethnicity: Any other Asian background	0.76	0.42
Ethnicity: Any other black background	-3.99 *	0.62
Ethnicity: Any other ethnic group	1.63 *	0.43
Ethnicity: Any other mixed background	0.85 *	0.40
Ethnicity: Any other white background	1.59 *	0.29
Ethnicity: Bangladeshi	1.21 *	0.42
Ethnicity: Black - African	-2.72 *	0.31
Ethnicity: Black Caribbean	-5.86 *	0.44
Ethnicity: Chinese	15.25 *	0.79
Ethnicity: Gypsy/Roma	-2.85 *	1.34
Ethnicity: Indian	5.28 *	0.33
Ethnicity: Information not yet obtained	1.21	0.72
Ethnicity: Pakistani	-0.89 *	0.30
Ethnicity: Refused	0.18	0.72
Ethnicity: Traveller of Irish Heritage	-3.04	4.24
Ethnicity: White - Irish	2.55 *	0.90
Ethnicity: White and Asian	3.70 *	0.51
Ethnicity: White and Black African	-1.49 *	0.70
Ethnicity: White and Black Caribbean	-3.41 *	0.45
Achieved grade 4 in English prior attainment	2.18 *	0.23
Achieved grade 4 in maths prior attainment	6.94 *	0.19
School or college in an urban area	1.18 *	0.21
Region: East Midlands	-2.14 *	0.24
Region: East of England	-0.62 *	0.21
Region: North East	2.52 *	0.28
Region: North West	4.11 *	0.21
Region: South East	-0.07	0.20
Region: South West	1.17 *	0.23
Region: West Midlands	-0.60 *	0.21
Region Yorkshire and the Humber	0.67 *	0.23
Main qualification type: GCSE	-77.64 *	0.37
Main qualification type: Other level 1&2	-76.16 *	0.19
Main qualification type: Non-academic level 3	-24.75 *	0.14
Institution type: Academy	2.23 *	0.19
Institution type: Colleges	-10.81 *	0.21

Dependent variable is best 3 point score, method 1 throughout	Coefficients and standard errors		
Institution type: Free schools	3.26	*	0.41
Institution type: Other	-12.92	*	2.98
Institution type: Sixth form college	2.20	*	0.23
Institution type: Special schools	-17.13	*	0.68
R squared	0.549		

* indicates significance at the 95% level

Next steps OLS model

Dependent variable is 16-19 attainment point score from 2007		
Independent variables	Coefficient	
Average of key stage 4 English and maths	-16.90	*
Average of key stage 4 English and maths squared	0.60	*
Difference between key stage 4 English and maths results	-1.80	*
Gender: female	17.10	
Free school meals (FSM)	-42.40	*
Future plans at 16: Further Education	5.90	
Future plans at 16: Higher Education	54.00	*
Future plans at 16: Unsure	7.70	
Parents' socio-economic class: mid	-26.40	*
Parents' socio-economic class: low	-8.90	
Parents' education: Level 2	-3.70	
Parents' education: Level 3	1.70	
Parents' education: Level 4/5	-9.50	
Parents' education: Level 6+	23.10	
Institution type: Academy school sixth form	10.20	
Institution type: FE college	32.90	*
Institution type: Sixth form college	74.50	*
Institution type: Tertiary college	7.20	
R squared	0.46	

* indicates significance at the 95 per cent level

Oaxaca-Blinder decomposition - Model results for disadvantaged and non-disadvantaged students

Dependent variable is best 3 point score, method 1	Coefficients and standard errors of models for disadvantaged and non-disadvantaged students				
Independent variables	Non-disadvantaged students		SE	Disadvantaged students		SE
(Intercept)	60.80	*	1.00	42.40	*	1.37
Gender: Female	3.70	*	0.11	3.29	*	0.19
Special Educational Needs (SEN)	0.75	*	0.20	0.14		0.25
English as an Additional Language (EAL)	9.31	*	0.63	0.96		0.77
Ethnicity: Any other Asian background	-4.51	*	0.44	1.50	*	0.72
Ethnicity: Any other black background	-2.61	*	0.79	0.89		0.79
Ethnicity: Any other ethnic group	-0.52		0.52	3.72	*	0.60
Ethnicity: Any other mixed background	-0.76		0.45	1.74	*	0.62
Ethnicity: Any other white background	0.98	*	0.31	2.78	*	0.52
Ethnicity: Bangladeshi	-2.91	*	0.52	2.53	*	0.57
Ethnicity: Black - African	-3.61	*	0.37	0.90	*	0.43
Ethnicity: Black Caribbean	-3.44	*	0.54	0.36		0.59
Ethnicity: Chinese	5.92	*	0.78	5.09	*	1.77
Ethnicity: Gypsy/Roma	-7.26	*	2.19	-3.76	*	1.63
Ethnicity: Indian	0.13		0.34	5.39	*	0.72
Ethnicity: Information not yet obtained	1.88	*	0.79	-0.85		1.19
Ethnicity: Pakistani	-2.47	*	0.33	2.07	*	0.45
Ethnicity: Refused	-1.56	*	0.79	2.33		1.20
Ethnicity: Traveller of Irish Heritage	-9.00		6.88	0.02		4.97
Ethnicity: White - Irish	0.50		0.91	-0.18		1.82
Ethnicity: White and Asian	-0.09		0.54	1.06		0.90
Ethnicity: White and Black African	-3.14	*	0.83	1.35		0.95
Ethnicity: White and Black Carribean	-2.86	*	0.54	0.83		0.62
Average of KS4 English and maths	-23.02	*	0.22	-12.40	*	0.26
Average of KS4 English and maths squared	3.61	*	0.02	2.31	*	0.03
KS4 English minus KS4 maths	-1.06	*	0.05	-1.00	*	0.07
EAL x KS4 English and maths	-1.45	*	0.10	0.23		0.14
School or college in an urban area	-0.13		0.21	0.33		0.49
Region: East Midlands	-0.32		0.26	0.61		0.42
Region: East of England	0.98	*	0.23	1.75	*	0.38

Region: North East	2.59	*	0.31	3.24	*	0.45
Region: North West	4.21	*	0.23	4.92	*	0.33
Region: South East	0.71	*	0.21	0.51		0.36
Region: South West	2.06	*	0.25	1.39	*	0.42
Region: West Midlands	-0.14		0.24	1.93	*	0.34
Region Yorkshire and the Humber	1.72	*	0.25	2.06	*	0.37
Average KS4 English and maths of cohort	6.90	*	0.12	5.86	*	0.18
Main qualification type: GCSE	-50.31	*	0.47	-46.44	*	0.58
Main qualification type: Other level 1&2	-48.16	*	0.24	-43.08	*	0.36
Main qualification type: Non-academic level 3	0.22		0.17	-0.41		0.30
Institution type: Academy	-0.12		0.20	-0.12		0.37
Institution type: Colleges	-0.58	*	0.26	-2.60	*	0.43
Institution type: Free schools	0.00		0.44	0.42		0.70
Institution type: Other	-7.03		3.78	2.64		5.21
Institution type: Sixth form college	2.50	*	0.24	0.09		0.43
Institution type: Special schools	8.74	*	2.46	4.01		2.27
R squared			0.603			0.596

* indicates significance at 95% level

Oaxaca-Blinder: Explained and unexplained components of disadvantage attainment gap

Dependent variable is best 3 point score, method 1	Variation between disadvantaged and non-disadvantaged student point scores, explained/unexplained components and standard errors			
Independent variables	Explained component of gap	SE	Unexplained component of gap	SE
Prior attainment variables	11.85	0.08	-17.51	0.70
EAL:Prior attainment interaction term	0.63	0.03	-2.16	0.18
Special Educational Needs	-0.06	0.01	0.10	0.04
Gender: Female	-0.02	0.01	0.21	0.11
English as an additional language (EAL)	-0.90	0.04	2.07	0.19
Region	0.02	0.02	0.01	0.04
Main qualification type	10.06	0.09	0.89	0.11
School or college in an urban area	0.00	0.01	-0.44	0.52
Average key stage 4 English and maths results of students at the same institution (cohort effect)	3.73	0.06	5.39	1.14
Institution type	0.23	0.04	1.42	0.90
Ethnicity	0.32	0.03	1.73	0.53
Constant			12.51	1.75
Total	25.88		4.21	

References

Department for Education and Skills, and UK Statistics Authority. 2004. *Statistics of Education: Variation in Pupil Progress 2003.*

Dominguez-Reig, Gerard, and David Robinson. 2019. '16-19 Education Funding: Trends and Implications'. Education Policy Institute. https://epi.org.uk/wp-content/uploads/2019/05/16-19-Funding_EPI-_2019.pdf.

Education and Skills Funding Agency. 2019. 'Funding Guidance for Young People 2019 to 2020. Funding Rates and Formula'. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/809266/Funding_rates_and_formula_201920_v2.pdf.

Evans, Helen. 2007. *Value-Added in English Schools*. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.178.9363&rep=rep1&type=pdf.

Hedges, Sophie, and Stefan Speckesser. 2017. 'Peer Effects and Social Influence in Post-16 Educational Choice'.

Hupkau, Claudia, Sandra McNally, Jenifer Ruiz-Valenzuela, and Guglielmo Ventura. 2017. 'Post-Compulsory Education in England: Choices and Implications'. *National Institute Economic Review*. https://doi.org/10.1177/002795011724000113.

Hutchinson, Jo, Mary Reader, and Avinash Akhal. 2020. 'Education in England: Annual Report 2020'. London: Education Policy Institute. https://epi.org.uk/wp-content/uploads/2020/08/EPI_2020_Annual_Report_.pdf.

'Key Stage 4 Performance 2019 (Revised)'. n.d. GOV.UK. Accessed 30 November 2020. https://www.gov.uk/government/statistics/key-stage-4-performance-2019-revised.

Maragkou, Konstantina. 2019. 'Socio-Economic Inequality and Academic Match among Post-Compulsory Education Participants'. Centre for Vocational Education Research. http://cver.lse.ac.uk/textonly/cver/pubs/cverdp023.pdf.

McIntosh, Steven. 2019. 'Post-16 Aspirations and Outcomes: Comparison of the LSYPE Cohorts', June, 19.

Patrignani, Pietro, Alice Battiston, and Gavan Conlon. 2019. 'BTECs, Higher Education and Labour Market Outcomes Using the Longitudinal Education Outcome (LEO) Dataset'. Centre for Vocational Education Research. http://cver.lse.ac.uk/textonly/cver/pubs/cverdp024.pdf.

Robinson, David. 2019. 'Further Education Pathways: Securing a Successful and Healthy Life after Education'. Education Policy Institute. https://epi.org.uk/publications-and-research/further-education-pathways/.

Thomson, Dave. 2018. 'Long-Term Outcomes: How Did Life Turn out for Those Who Took Level 3 BTEC Qualifications?' *FFT Education Datalab* (blog). 20 November 2018. https://ffteducationdatalab.org.uk/2018/11/long-term-outcomes-how-did-life-turn-out-for-those-who-took-level-3-btec-qualifications/.

Wiliam, Dylan. 2010. 'Standardized Testing and School Accountability'. *Educational Psychologist* 45 (2): 107–22. https://doi.org/10.1080/00461521003703060.