Technical appendix

Education in England: Annual Report 2020 examines trends in overall attainment and attainment gaps for the period 2011 to 2019. With the exception of our post-16 destinations analysis, all results are drawn from the National Pupil Database (NPD).[1]

Since our 2018 report, DfE has changed its data access arrangements and as a result we have accessed new versions of the data files used to produce this report. The new files are not identical to the previous versions and we have updated all our time series from the new files. This has resulted in some small changes to figures for historical years. Therefore, it is recommended that figures in this report are *not* compared with those previously published in order to draw any conclusions about trends. Please refer to the time series within this year's report in order to look at trends.

Pupil population

We include pupils at all state-funded schools except for those whose sole, or main, registration was in alternative provision, a pupil referral unit, or a hospital school. Independent schools are not included, apart from a small number of providers in the Early Years Foundation Stage.[2]

Pupil attainment

Our report uses point scores rather than threshold measures to assess system performance.

Early Years Foundation Stage

For the early years, we use total points score in the Early Years Foundation Stage Profile (EYFSP) to measure attainment. The EYFSP is a teacher-assessed measure of pupil proficiency across seventeen learning goals, with children assessed as either meeting the level of development expected (score=2) at the end of the reception year, exceeding this (score=3), or not having reached the level (score=1). The total points score aggregates scores across the seventeen goals, ranging from a minimum score of 17 to a maximum of 51.

We use the total point score, as opposed to the proportion of children achieving the Government's benchmark of "a Good Level of Development". Measuring performance using a binary benchmark can lead practitioners to, consciously or otherwise, push pupils over the line. Our use of the total point score reduces the effect of this – thereby giving a more reliable picture of the trend in attainment.

Note that as the EYFSP scoring system changed in 2013 and the score distributions differ significantly due to reductions in curriculum breadth, we only display results from 2013 to 2019.

Key Stage 2

At Key Stage 2 (KS2), attainment is measured using the average of reading and mathematics scaled scores. Scaled scores for these domains are derived from national test results and can take values between 80 and 120. We also include teacher-assessed attainment scores for pupils who do not reach the lowest measurable score in the test, whose scores range from 59 to 79, to enable them to be included in the point distribution. Where pupils are missing either result, the average takes the

---

[1] www.gov.uk/government/collections/national-pupil-database
[2] E.g., in 2017, 246 children in the EYFSP cohort were registered across 35 independent and non-maintained schools.

value of the subject they do have a score for. Where neither subject has a score, the pupil is not included in our analysis.

It is important for our analysis to use the most consistent measures possible over time. We therefore do not draw on the spelling, punctuation and grammar assessment as it was only introduced in 2013, nor the assessment in writing as from 2012 it has been teacher-assessed and was not comparable with tests taken in prior years.

2016 was the first year pupils were assessed against a new national curriculum, in tests that were designed to be more difficult, and with a new scoring system. These changes make it impossible to make direct comparisons between the 2016 results and years prior to then for assessing overall attainment levels. The material assessed and range of test scores does support comparison of relative attainment differences between groups ('gaps') before and after 2016, however.

Key Stage 4

We use average GCSE grade per subject to measure Key Stage 4 (KS4) attainment. While this measure excludes non-GCSE qualifications, it does include AS level qualifications completed in KS4. Scores range from 0 to 10.75.[3]

Recent reforms to GCSEs make it difficult to compare performance over time. By using an average point score (as opposed to total points), it makes this comparison more consistent when we come to assess the disadvantage gap, because it is not affected by the number of GCSEs taken, but takes into account GCSE subjects across the curriculum. Holding constant the qualification types included in the measure by restricting this to GCSEs is helpful for assessing changes in the gap before and after 2014, when major reductions were made to the range of qualifications schools could use to meet performance floor standards.

The average GCSE grade per subject measure is provided in the NPD for all the years we analyse. To account for changes in the point scores grades awarded in 2016, 2017 and again in 2018 (for unreformed GCSEs), we adjust average scores in prior years by mapping across the old score boundaries to the new, and interpolating to produce an adjusted figure. We make no adjustment for the introduction of the new nine grade scale (rather than eight) in 'reformed' GCSE English and maths in 2017, further subjects in 2018, and the final remaining subjects in 2019.[4] The gradual shift from unreformed to reformed GCSEs makes strict comparisons of attainment results impossible, but this does not affect our ability to make disadvantage gap comparisons over time, as we are effectively measuring the change of within-year rank of various pupil groups, not absolute scores.[5]

We use average GCSE grade per subject as, unlike other measures (e.g. a broader measure of attainment that includes non-GCSE subjects in KS4, or Attainment 8 points), it has not been affected

---

[3] AS level, reformed GCSE and unreformed GCSE subjects count towards this measure, and each have their own grading scale and associated point scores. An average above 8.5 (the maximum score for unreformed GCSE subjects) is rare, accounting for fewer than 0.02 per cent of KS4 pupils in 2017. For information on the point score scales for contributing subjects, see: www.gov.uk/government/publications/key-stage-4-qualifications-discount-codes-and-point-scores

[4] https://www.gov.uk/government/publications/get-the-facts-gcse-and-a-level-reform/get-the-facts-gcse-reform

[5] For details of how the resilience of the ranking approach to grading system changes has been tested, please see:
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/398657/SFR_40_2014_Measuring_disadvantaged_pupils_attainment_gaps_over_time__updated_.pdf

by changes in which non-GCSE qualifications count in Department for Education's (DfE) school performance tables.[6]

To provide maximum stability in our headline time series, when measuring the disadvantage gap we also calculate the gap based on the average of GCSE English and maths. This provides a measure that, while quite narrow, is not affected by changes in GCSE subject entry patterns.

Attainment gaps

We continue to report attainment gaps between specific pupil groups.[7] We calculate these gaps using the same mathematical procedure as the DfE, though we present our results in 'months of progress' terms, and apply these calculations to different attainment measure inputs.[8] The steps followed to calculate the gap are:

1. Rank all pupils by score, as per the attainment measures described earlier.

2. Identify the relevant groups of interest, and calculate the mean rank of pupils in these groups.

3. Subtract the rank of the group of interest from that of the reference group used.

4. Convert this rank difference to a months of progress measure, using a multiplier of 33 for the Early Years Foundation Stage Profile, of 64 for Key Stage 2, and of 99 for Key Stage 4.[9]

Pupil characteristics

Disadvantaged

We define disadvantaged pupils as those who have been eligible for free school meals (FSM) in any of the prior six years. This is used as the deprivation eligibility criterion for the Pupil Premium. The reference group these pupils are ranked against is those who have never received FSM over the last six years.

For early years, we do not have a six-year history of FSM eligibility, so instead we measure disadvantage by whether they are eligible for FSM in the current academic year.

Persistently disadvantaged

We define persistently disadvantaged pupils as those who have been eligible for free school meals (FSM) for 80 per cent or more of their time in school. We include as much of pupil's school life as possible, with the constraint that school census data started in 2002 and we need to keep the period constant across cohorts. We therefore go back to Year 2, since for the KS4 cohort in the first year in our time series – 2011 – the earliest available data are from when they were in Year 2.

---

[6] The 2011 Wolf Review recommended the removal of many non-GCSE qualifications from school performance tables, which saw over 3,000 of these qualifications removed in 2014.

[7] J. Andrews, J. Hutchinson and D. Robinson 'Closing the Gap? Trends in Educational Attainment and Disadvantage', August 2017

[8] The DfE methodology for calculating attainment gaps for disadvantaged pupils can be found here: www.gov.uk/government/statistics/measuring-disadvantaged-pupils-attainment-gaps-over-time

[9] This updates the methodology used for calculating gaps in our 2016 annual report. For further discussion of that methodology see 'Education in England: Progress and goals' https://epi.org.uk/report/ambitions-forenglish-education/. We have applied the mean rank approach to earlier years to calibrate difference in ranks against months of progress.

For KS4, this results in data from the ten-year period prior to assessment in Year 11. For example, for the KS4 2019 cohort, we use FSM data from 2010 inclusive; for the 2011 KS4 cohort, we use FSM data from 2002 inclusive.

For KS2, we are able to use data from the five-year period prior to assessment in Year 6. For example, for the KS2 2019 cohort, we use FSM data from 2015 inclusive; for the KS2 2007 cohort, we use FSM data from 2002 inclusive. Note that unlike our KS4 calculations which start in 2011, we have a longer time series for KS2 that stretches back to 2007.

The mean rank of persistently disadvantaged pupils is compared with the mean rank of pupils who have neither met the definition of disadvantage in the previous paragraph, nor met the definition of persistent disadvantage (i.e. been eligible for FSM for 80 per cent or more of their school life). We use this reference group so that we use the same comparison group for the disadvantage gap and the persistent disadvantage gap.

Ethnicity

For this characteristic, we express the gap for all ethnic groupings relative to White British pupils (who are the largest group, representing 69 per cent of the KS4 pupil population in 2019).

Special Education Needs and Disabilities (SEND)

We report the gap for two SEND categories:

- pupils with a statement of special educational needs or an education, health and care plan (EHCP); and
- pupils with an identified special educational need but without a statement or plan.

Both are reported relative to pupils with no identified SEND.

Late arriving English as an additional language (EAL)

We define late arriving EAL pupils as those who are recorded as having EAL, and who have entered the English state-school system in either Year 5 or Year 6 for key stage 2, or either Year 10 or Year 11 for key stage 4. The reference group against which these pupils are ranked is the group of pupils who are recorded with English as their first language in the current year, and who have never in the past been recorded as having EAL.

We do not report the attainment gap for pupils who fall outside of these two groups (e.g., those who are EAL, but appeared in the state school system prior to the last two years). This is because the total EAL group is so heterogenous in terms of English language proficiency and attainment that average attainment measures tend to be meaningless in representing individuals within the group.

Geographic breakdowns

We also report the gap on a geographic basis, covering local authorities (LAs), parliamentary constituencies, opportunity areas, regional school commissioner regions and city regions. In each we construct the gap by ranking the (persistently) disadvantaged pupils in the area relative to the national mean rank of those who not (persistently) disadvantaged. We do this rather than express the rank in terms of the difference between disadvantaged and non-disadvantaged pupils *within* the area to allow for a consistent reference point across areas. This avoids representing disadvantage gaps as being especially large in certain geographic areas based on very high attainment of non-disadvantaged children in the area, rather than low attainment by disadvantaged children.

We classify geographical breakdowns based on pupil residence instead of the location of the school they attend. We do this because local authorities are not accountable for all schools within their area. This makes attainment more comparable across phases and between local authorities, as the geographical breakdowns are not influenced by differential secondary school admissions policies which can result in transfers of pupils across LA boundaries, thereby risking the introduction of bias into our estimates of the disadvantage gap.

Analysis by parliamentary constituency

Pupils are allocated to parliamentary constituencies based on their home address. Note that only pupils who are resident in England are included in the analysis. In a small number of cases, pupils may travel across a national border (e.g. from Wales or Scotland to England) to go to school; these cases are excluded.

Our pupil-level data provide residence information at the Lower Layer Super Output Area (LSOA) geographic level and not at parliamentary constituency level. As LSOAs are a measure of *census* geography, and parliamentary constituencies of *electoral* geography, the two do not always nest conveniently together[10]. A best-fit methodology was therefore developed to derive parliamentary constituencies from LSOA data:

- In cases in which LSOAs nest completely within parliamentary constituencies, these are assigned by default.[11]
- In cases in which an LSOA is split between two parliamentary constituencies, a majority rule is applied.
    - If **more than half** of an LSOA's component output areas is within a given parliamentary constituency, all pupils within that LSOA are allocated to that constituency[12].
    - If **exactly half** of an LSOA's component output areas are within two different parliamentary constituencies, all pupils within that LSOA are allocated to one of these two constituencies by random selection[13].

Office for National Statistics geography lookups[14] were used to first match output areas to LSOAs, and then to parliamentary constituencies.

Gap by GCSE subject

For the first time, in this year's Annual Report we include estimates of the disadvantage gap by GCSE subject for 2019. We include all subjects for which there is attainment data in the NPD in this academic year.

---

[10] For more information on UK geographies, see the Office for National Statistics (ONS) guidance available here: https://www.ons.gov.uk/methodology/geography/ukgeographies.
[11] This was the case for > 99% of LSOAs.
[12] In the 2019 data, this was the case for < 1% of LSOAs.
[13] In the 2019 data, this was the case for < 0.1% of LSOAs.
[14] Available here: http://geoportal.statistics.gov.uk/.

For subjects which have not yet converted to the new 9-1 grading system (Portuguese, Gujarati, Turkish, Persian and Biblical Hebrew), we convert scores manually in order to maintain consistency across subjects.

In Figure 2.5, a bubble is shown for 'all languages'. This aggregates attainment across language subjects, many of which have very small pupil intakes, by taking the best score achieved by a given pupil in their language subject(s).

Analysis of the impact of persistent disadvantage on the gap

This year we conduct an analysis of the impact of persistent disadvantage on the gap. This has two parts. First, we look at the impact of changes in persistent disadvantage at the national level on the evolution of the gap over time. Second, we look at the impact of differential levels of persistent disadvantage across regions on the gap by region in 2019.

Analysis by year

As Figure 3.2 in the report details, the level of persistent disadvantage experienced by disadvantaged pupils has changed over the last decade. Over the last five years, persistence has increased for disadvantaged pupils, whereas prior to that it was decreasing. As this broadly mirrors changes in the KS4 disadvantage gap at national level, there is a possibility that changes in the disadvantage gap are being driven by these increases in persistent poverty.

First, we adopted a simple regression approach to look at the impact of four measures of persistent disadvantage on the disadvantage gap (in GCSE English and maths) at national level. Despite the small sample size (9 data points), we found that changes in persistent disadvantage *among the disadvantaged group* were a statistically significant predictor of the disadvantage gap. By contrast, changes in persistent disadvantage across the whole pupil population were not significant.

The following table summarises the results of our year-level regression analysis. Model 1 measures persistence as the mean percentage of pupils' school lives that they are eligible for FSM, across the whole pupil population. Model 3 uses the same definition, but just among disadvantaged pupils. Model 2 measures persistence as the proportion of pupils who met the definition for persistent disadvantaged – that is, they are eligible for FSM for 80 per cent or more of their school lives. Model 4 uses the same definition, but just among disadvantaged pupils.

	Model 1	Model 2	Model 3	Model 4
Coefficient	25.48	31.13 ***	1.47	0.26 ***
Standard error	(111.97)	(4.82)	(0.63)	(0.03)
R-squared	0.007	0.8565	0.4403	0.9233

These results suggest that it is changes in persistence among the disadvantaged group that are salient in terms of the national evolution of the gap over time.

Given this, we divide the disadvantaged group into five subgroups based on how persistently disadvantaged they are:

- 0-19 per cent of their school life (low persistence)
- 20-39 per cent of their school life (low-medium persistence)
- 40-59 per cent of their school life (medium persistence)
- 60-79 per cent of their school life (medium-high persistence)

- 80-100 per cent of their school life of their school life (high persistence, i.e. persistently disadvantaged).

This enables us to show the impact of changes in persistence in a more intuitive way, as well as to investigate non-linear time trends.

Analysis by region

As Gorard et al. (2019) discuss, comparing raw disadvantage gaps across regions without accounting for differential levels of persistent poverty across regions is likely to give a biased impression of the relative performance of regions.

While we are not able to do this for regional school commissioner areas, opportunity areas or and city regions due to small sample sizes, at local authority and parliamentary constituency levels, we employ a regression approach and control for different measures of persistent disadvantage to see what impact it has on the gap. Our regression model is the following:

$$\begin{aligned} Raw\ disadvantage\ gap \\ = (\beta * Mean\ persistent\ disadvantage) + \ Adjusted\ disadvantage\ gap \end{aligned}$$

In other words, by regressing the mean level of persistent disadvantage on the disadvantage gap at local authority and parliamentary constituency level, we are able to 'net out' the impact of persistence on the gap and reach an estimate of the residual or 'adjusted' disadvantage gap.

In order to show how the gap consequently increases for some areas and decreases in others, we apply the coefficient from the regression to the national mean level of persistence. This enables us to calculate an adjusted gap for each area which represents what the gap would be like *if that area had the national level of persistence.*

In the local authority level regressions, we found that London confounded our results in such a way that the findings did not make sense. This is because while London represents a relatively small share of the total pupil population in England, it is overrepresented in terms of the number of local authorities, with a third of English LAs being in London. While the drivers of the 'London effect' – the relative overperformance of schools in London given its poverty levels – are contested, it is widely acknowledged that its ethnic diversity plays a significant role.[15] To control for the London effect, we therefore use ethnicity as a proxy to help us better approximate the impact of persistent disadvantage. We do this by including dummy variables for each ethnic group by region as controls. The results from this analysis are more sensical and they corroborate our findings at parliamentary constituency level. Note that we use ethnicity here to better approximate the relationship between persistent disadvantage and the gap; we adjust for persistence by extracting the persistence coefficient from these multivariate regressions without adjusting for ethnicity itself.

Table 2 and Table 3 show our regression results at local authority and parliamentary constituency level respectively. We use Model 1 to calculate the adjusted gaps, on the basis that it has the highest R-squared at both local authority and parliamentary constituency level.

---

[15] Burgess, S. (2014) 'Understanding the success of London's schools'. Working paper 14/133, The Centre for Market and Public Organisation, Bristol University.

Table 2: Regression results at local authority level

	Model 1	Model 2	Model 3	Model 4
Persistence coefficient	20.1 **	18.87 ***	25.58 ***	13.69 ***
Standard error	(4.28)	(4.89)	(5.79)	(3.73)
R-squared	0.79	0.78	0.79	0.78
Ethnicity controls	Yes	Yes	Yes	Yes

Table 3: Regression results at parliamentary constituency level

	Model 1	Model 2	Model 3	Model 4
Persistence coefficient	25.8 ***	23.29 ***	33.82 ***	15.93 ***
Standard error	(2.66)	(2.83)	(3.66)	(2.18)
R-squared	0.61	0.59	0.60	0.58
Ethnicity controls	Yes	Yes	Yes	Yes

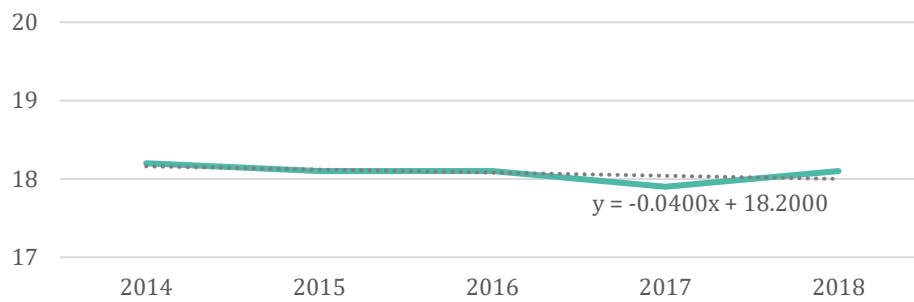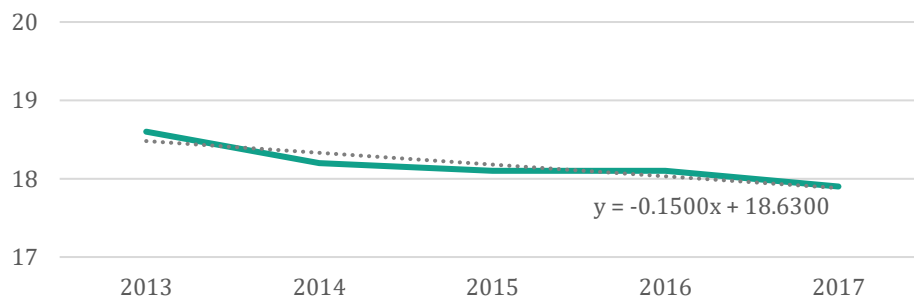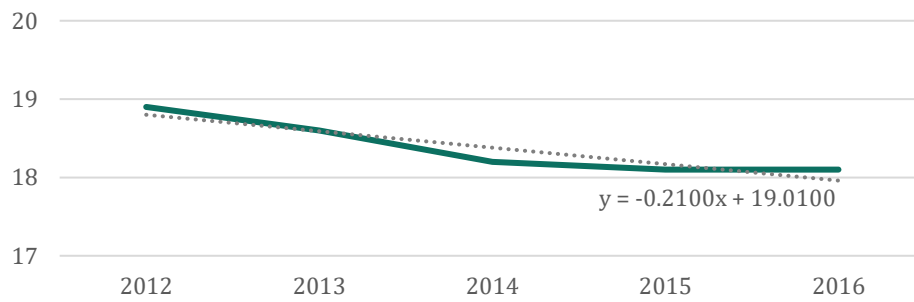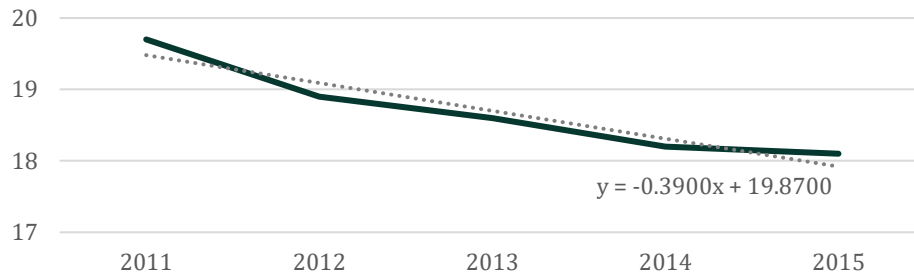## Key Stage 4 disadvantage gap projections

For the disadvantage gap, we create a yearly projection of how long it will take the gap for a given key stage to close, based on the most recent five years of data. A linear trendline is fitted to each five-year trend and the equation of this trend line is used to calculate the number of years until the gap reaches zero. This procedure was repeated for five five-year periods, as follows:

- 2011 - 2015
- 2012 - 2016
- 2013 - 2017
- 2014 - 2018
- 2015 - 2019

These trends are shown below. The vertical axis of the charts represents the difference in the percentile ranks of disadvantaged versus other pupils, based on their attainment in GCSE English and maths.

The change in the number of years until the gap is expected to close is used to determine whether the gap is closing more quickly or more slowly in recent years than in prior years. It can be visually seen by changes in the gradient or 'steepness' of the trend line across the five five-year periods ending in 2015, 2016, 2017, 2018 and 2019.

5-year trends in the GCSE English & maths disadvantage gap (months)

y = -0.3900x + 19.8700

y = -0.2100x + 19.0100

y = -0.1500x + 18.6300

y = -0.0400x + 18.2000

y = -0.0000x + 18.0600

9
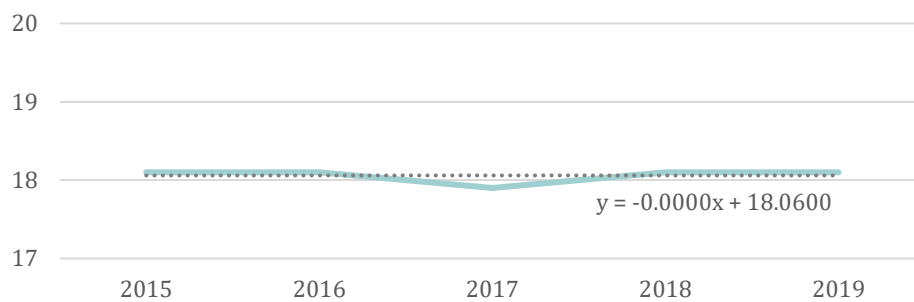
Change in gap for LAs with similar gaps in 2012

To illustrate how LA attainment gaps have shifted over time, we compare relative changes in the size of gap from 2012 to 2019.[16] We first regress present 2019 gaps on 2012 gaps for all LAs. For each LA, we then use this regression model to estimate the 2019 gap (in effect, an estimate of the gap compared to LAs that had similar gaps in 2012). The change in the gap shown is the difference between each LA's estimated 2019 gap and their actual gap (positive figures indicating the LA has narrowed more than the estimate, and vice versa). This approach is used as the actual change in gap (2019 gap minus 2012 gap) is heavily correlated with the size of the 2012 gap, and therefore tells us little about relative local authority performance in narrowing the gap.

Destination gaps at post-16

To measure segregation in the post-16 destinations of KS4 pupils, we use an 'index of dissimilarity'.[17] This index has been applied to measure several forms of segregation in educational contexts previously.[18]

Our analysis uses Key Stage 4 education destinations data from DfE. In the latest data release, DfE have used an updated methodology which has been applied to 2017 data and all prior years. Changes predominantly refer to how double-counting is dealt with and how some destinations are recorded.[19] We have updated our segregation index time series to reflect the new DfE methodology.

Our dissimilarity index measures segregation across the following post-KS4 destinations:

- Further education (FE) college or other FE
- 6th form: college or secondary school
- Other education destination (e.g. includes special schools, independent schools, alternative provision, higher education institutions, and post-16 specialist institutions)
- Apprenticeships
- Sustained employment and/or training destination[20]
- Destination not sustained (e.g. those who participated in education or employment for less than two terms, or who had no participation and claimed out-of-work benefits).

Simply put, the index measures how evenly two groups (in our case disadvantaged and non-disadvantaged pupils) are distributed across destinations relative to their share of the total pupil population. For example, if England has a 1:10 ratio for disadvantaged pupils to all other pupils, then the index will quantify how far away we are from achieving a 1:10 ratio in each post-16 destination.

---

[16] For the early years we use 2013 rather than 2012 as our baseline year, as this was the first year to use the EYFSP.

[17] This is the only piece of analysis that does not draw on the NPD. Instead we use DfE Destinations data available here: https://www.gov.uk/government/collections/statistics-destinations

[18] See Allen, B. & Vignoles, A. (2006) *What Should an Index of School Segregation Measure?* Centre for the Economics of Education, London School of Economics and Political Science. Available: http://cee.lse.ac.uk/ceedps/ceedp60.pdf; and Whitehurst, G.J., Reeves, R.V., & Rodrigue, E. (2016) *Segregation, Race and Charter Schools: What do we know?* Brookings Institute. Available: https://www.brookings.edu/wp-content/uploads/2016/10/ccf_20161021segregation_version-10_211.pdf

[19] For further details, see the destinations methodology: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/748208/Destinations_Quality_Methodology_Information_2017.pdf

[20] To count as a 'sustained' destination, the young person has to be participating for at least 'two terms' or 'six months' of the academic year after they have completed Key Stage 4.

The formula used to generate the index is:

$$S = \frac{1}{2} \sum_{i=1}^{N} \left| \frac{d_i}{D} - \frac{a_i}{A} \right| \; x \; 100$$

S = Dissimilarity index

d_i = number of disadvantage pupils in destination i

D = total population of disadvantage pupils

a_i = all other pupils in destination i

A = total population of all other pupils

The index can take a value between 0 and 100. 0 indicates a complete absence of segregation and 100 indicates total segregation. The result can be interpreted as the proportion of the disadvantaged cohort who would need to change destinations to achieve an absence of segregation.

It is important to note that this measure does not assume any hierarchy of post-16 destinations in terms of desirability. It simply quantifies how alike the post-16 trajectories of disadvantaged pupils and their peers are.