Covid-19 disruptions
Attainment gaps and primary school responses

Report

May 2021

Ben Weidmann, Rebecca Allen, Dave Bibby, Rob Coe, Laura James, Natasha Plaister and Dave Thomson

Please cite as:

Weidmann, B., Allen, R., Bibby, D., Coe, R., James, L., Plaister, N. and Thomson, D., Covid-19 disruptions: Attainment gaps and primary school responses, Education Endowment Foundation, 2021

The Education Endowment Foundation (EEF) is an independent grant-making charity dedicated to breaking the link between family income and educational achievement, ensuring that children from all backgrounds can fulfil their potential and make the most of their talents.

The EEF aims to raise the attainment of children facing disadvantage by:

- identifying promising educational innovations that address the needs of disadvantaged children in primary and secondary schools in England;
- evaluating these innovations to extend and secure the evidence on what works and what can be made to work at scale; and
- encouraging schools, government, charities, and others to apply evidence and adopt innovations found to be effective.

The EEF was established in 2011 by the Sutton Trust as lead charity in partnership with Impetus Trust (now part of Impetus: Private Equity Foundation) and received a founding £125m grant from the Department for Education. Together, the EEF and Sutton Trust are the government-designated What Works Centre for improving education outcomes for school-aged children.

For more information about the EEF or this report please contact:

Jonathan Kay
Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
SW1P 4QP

0207 802 1653

jonathan.kay@eefoundation.org.uk

www.educationendowmentfoundation.org.uk

Contents

About the evaluator

This project was run by colleagues at FFT, TeacherTapp, and the EEF. The team consisted of Ben Weidmann, Dave Thomson, Natasha Plaister, Becky Allen, Rob Coe, Dave Bibby, and Laura James.

Acknowledgements

We would like to thank the school support team at FFT for their efforts to keep strong relationships with schools in collecting both assessment and teacher survey data, which enabled analysis to go ahead. We also thank Jennifer Stevenson, Celeste Cheung, and Jamila Boughelaf at the EEF for their support and guidance throughout the project.

Executive summary

The study

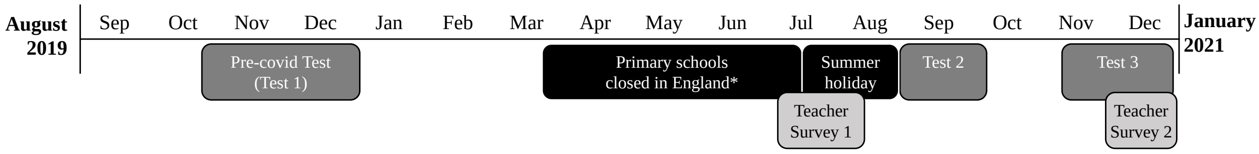
This descriptive study focuses on primary schools in England. It has three aims:

1. **Quantify changes in attainment gaps since the onset of covid-19**. Our focus is on gaps between economically disadvantaged pupils[1] and their peers in Years 2 to 6.
2. **Describe how teachers and schools responded to the challenges of covid-19.** We examine two periods: first, the 'initial closure period' in March–May 2020 when schools were suddenly required to shift to remote learning; second, the 're-opening period' when pupils returned to schools for the start of the 2020–2021 academic year.
3. **Explore associations between school responses and changes in attainment gaps**. We explore associations between the way in which schools responded (during the initial lockdown and during the Autumn 2020 term) and changes in attainment gaps.

This is an observational study. The sample sizes differ across research questions and range between 8,500 and 13,500 pupils (in 65 to 97 schools). Our samples are similar to the broader population in terms of prior achievement and levels of disadvantage. The study uses data from three assessment points (for reading and maths). At all three points, the same sample of students sit standardised RS Assessment tests with the first assessment completed in 2019 prior to the covid pandemic. The study also administered two teacher surveys. The first survey focused on the initial period of lockdown; the second focused on the autumn 2020 term, when students returned to face-to-face instruction.

This is an interim report. The study will continue until the end of this school year.

Table 1: Summary of study findings



Note: *From 1 June 2020 schools were re-opened to children in Reception, Year 1, and Year 6 but remained closed to a vast majority of students in other year groups.

Research question	Finding
How did the attainment gaps between 'free school meal status in the past 6 years' (FSM6) pupils and their peers change between Autumn 2019 and Autumn 2020?	We find evidence that the attainment gaps between free school meal (FSM)-eligible pupils and their peers for primary maths have widened since Autumn 2019. On average, for Years 2–6, the maths gap between disadvantaged pupils and their peers widened by an estimated 0.07 effect size (ES) units (with a plausible range of 0.04 to 0.11). This represents an increase of between 10% to 24% of the pre-covid attainment gap. There was no discernible change in the disadvantage gap for reading.
How did 'attainment gaps' change during the Autumn 2020 term, when most pupils returned to 'face-to-face' schooling?	There were no clear changes in attainment gaps between FSM-eligible pupils and their peers in primary maths or reading, during the Autumn term in 2020. The attainment gap for maths was estimated to change by −0.01 ES units, with a range of [−0.04, 0.02]. For reading, the change was 0.01 ES units, with a range of [−0.02, 0.04]. Note that negative numbers indicate a widening of gaps.

[1] Defined by free school meals (FSM) status in the previous 6 years.

Research question	Finding
What were the associations between how schools responded to ‘remote learning’ from March–May 2020 and changes in attainment gaps between FSM6 pupils and their peers?	We examined five variables that describe primary schools’ responses to remote learning in March–May 2020: ‘phoning students’; ‘timetabling’; ‘live or recorded lessons’; ‘frequency of work submission’; ‘use of technology platforms’. Broadly, we found no evidence of clear associations between these variables and changes in attainment gaps.
What were the associations between school responses during the re-opening period (September–December 2020) and changes in the attainment gaps between FSM6 pupils and their peers in December 2020?	We examined five variables that describe primary school practice in the Autumn 2020 term: ‘providing videos/live streams for absent *pupils*’; ‘providing videos/live streams for absent *classes*’; ‘extra learning time’; ‘reducing the curriculum’; and ‘more small group interventions’. Overall, we found limited evidence of clear associations between these variables and changes in attainment gaps. However, there was some tentative evidence that providing video/live lessons to pupils who are absent is associated with reductions in attainment gaps.

Additional findings

- Years 2 and 3 in maths saw the largest estimates for gap widening. While this matches initial findings from other studies, suggesting that younger children have been worse affected by covid-19 disruptions (Curriculum Associates, 2020; Juniper Education, 2020), we do not have enough data from each year group to make firm conclusions about whether attainment gaps have grown more for younger children in maths, relative to other subject areas or age groups.

- There was considerable variation in how schools responded to the initial lockdown (March–May 2020). For example, 23% of teachers reported phoning students at least once a week, while 37% never phoned students (and 40% phoned students once or twice during the half-term). Similarly, we observed a variety of different approaches in how schools dealt with lost learning and covid-related disruptions when in-person schooling resumed in Autumn 2020. For example, 24% of schools provided live or recorded video lessons for individual students who were absent during this period, while 76% did not.

- Our analysis of the association between school responses and changes in attainment gaps had several limitations. In particular: we collected no qualitative information about how different responses were implemented. Moreover, we could only analyse data at the school level. This level captured some, but certainly not all, of the variation in teacher responses (27% to 73% depending on the variable).

- We do not find clear evidence of associations between school approaches to remote learning from March–May 2020 and relative changes in attainment from Autumn 2019 to September 2020. In other words, there were no clear associations between changes in attainment and: ‘phoning students’; ‘timetabling’; ‘live or recorded lessons’; ‘frequency of work submission’; ‘use of technology platforms’.

- We do not find clear evidence of associations between variables describing how schools approached in-person schooling in Autumn 2020 and relative changes in attainment from September 2020 to November/ December 2020. We saw no clear associations between changes in reading/maths attainment and: ‘providing videos/live streams for absent pupils’; ‘providing videos/live streams for absent classes’; ‘extra learning time’; ‘reducing the curriculum’; and ‘more small group interventions’.

Introduction

Background and motivation

This section provides the background and motivation to our study, which has three parts.

- In part 1, we **quantify changes in attainment gaps since the onset of covid-19**. We focus on gaps between economically disadvantaged pupils and their peers. As our measure of economic disadvantage is 'free school meal status in the past 6 years' (FSM6), we interchangeably refer to these inequalities in educational attainment as 'disadvantage gaps', 'FSM gaps', or 'economic gaps'.
- In part 2, we **describe how teachers and schools responded to covid-19 disruptions**. There are two periods of interest. First, the 'initial closure period', from March to May 2020. Second, the 're-opening period', from September 2020 to December 2020, when most students returned to school. What strategies and remedies did teachers adopt?
- Part 3 combines the first two. We **explore associations between school responses** (as captured in teacher surveys) **and changes in disadvantage gaps in attainment**. Again, we examine the 'initial closure period' and the 're-opening period'.

1. Measuring changes in educational inequalities during covid-19

Covid-19 caused unprecedented disruptions to schooling. Existing research shows how these disruptions have exacerbated inequalities in access to education. During the first period of school closures (March–May 2020) data from England suggests that, compared to their more affluent peers, children in low income households spent less time on education, had less parental support in terms of time and expertise, received fewer paid-for educational services, and had more problems with access to devices and the internet (Andrew *et al.*, 2020; Cullinane and Montacute, 2020; Hupkau and Petrongolo, 2020). Research from the United States shows how students in economically disadvantaged households engaged less with online educational platforms (Bacher-Hicks *et al.*, 2020; Chetty *et al.*, 2020).

A growing set of studies has focused on how these inequalities in educational inputs translate into attainment gaps. We build on this literature by estimating changes in attainment gaps between disadvantaged students and their peers in Years 2–6. Existing literature has tended to show that gaps in attainment have widened since the onset of the pandemic. We review specific findings below. First, however, we note that the literature has faced several challenges in providing estimates of how covid-19 has impacted economic gaps in attainment. These challenges include:

- **Making comparisons across time (pre- and post-covid) using a longitudinal sample or comparable samples.** Covid disruptions have led to large changes in the composition of students who take assessments, compared to those who do not sit tests. Many in-person tests have been cancelled or heavily disrupted. Attrition from online assessments has also been high, and may differ systematically across socioeconomic groups (Domingue *et al.*, 2021). In making comparisons across time—especially in examining whether (and by how much) disadvantage gaps in attainment are changing—it is important to compare apples with apples. The gap between FSM pupils and their peers can be influenced by the composition of the sample at different time points.
- **Using standardised measures of attainment, administered in a consistent way across time.** Ideally, attainment data should be standardised, rather than based on teacher assessments. To get a clear picture of how the magnitude of underlying inequalities may be changing, tests should be conducted in the same way (preferably in person) across different time periods. If, for example, testing conditions after the onset of covid-19 became systematically more difficult for economically disadvantaged students, which is especially plausible for tests administered at home, then resulting estimates of changes in gaps may not reflect underlying differences in learning.
- **Using individual-level indicators of disadvantage.** Quantifying attainment gaps between disadvantaged pupils and their peers using school-level or regional-level measures of disadvantage provides indirect estimates. In many cases these estimates will understate the magnitude of gaps. As an example, consider an analysis comparing the attainment of pupils in 'low-FSM' schools (defined as those with a below-average percentage of FSM pupils) with attainment in 'high-FSM' schools (defined as having

an above-average percentage of FSM pupils). In the 'high-FSM' schools, it may be the case that a majority of the pupils may not be from disadvantaged backgrounds. This will tend to mask any gaps.

- **Providing uncertainty estimates for how gaps have changed.** As most system-wide tests have been cancelled (eliminating the possibility of population-level estimates), efforts to quantify the way in which economic gaps have changed are subject to sampling uncertainty. In some analyses, uncertainties are not reported. This is often the case when analyses are focused on 'learning loss' and not on directly estimating attainment gaps. It is also important to distinguish between estimating an attainment gap and estimating a change in that gap before and after covid-19. Disadvantage gaps will generally have existed before covid-19 and the change may be small relative to the precision of a sample estimate.

Our study, described below, aims to address these difficulties. We use a longitudinal sample, in which the same students sat standardised, in-person tests at three time-points: November/December 2019 (pre-covid); September 2020 (when schools returned to in-person instruction); and December 2020 (after most children had been back at school for a full term). Our analytical approach explicitly focuses on estimating gaps and the associated uncertainties. Using these data, our first contribution is to provide direct estimates of changes in the disadvantage gap for primary reading and maths (Years 2–6). With these challenges in mind, we now turn to reviewing existing estimates.

Existing estimates of changes in attainment gaps due to economic disadvantage

Two studies in England use standardised assessments to provide direct evidence about changes in attainment gaps due to economic disadvantage.[2] RS Assessment (2020) estimate disadvantage gaps in reading and maths for Years 1–6 for Autumn 2019 and Autumn 2020. The report finds that, averaging across Years 1–6, gaps widened by 0.05 ES units for reading and 0.06 for maths.[3] A second study, by NFER, estimated the FSM gap in Year 2 for reading and maths in Autumn 2020, but was unable to calculate a comparable figure for the pre-covid period due to lack of information about pupils' FSM status (NFER, 2021).[4] These studies both conclude that economic gaps grew from Autumn 2019 to Autumn 2020.

Moving beyond the UK, several studies provide direct evidence about the extent to which economic gaps have grown. For example, Pier *et al.* (2021) use a similar approach to DfE (2021) to analyse data in California. They report evidence of widening gaps in elementary/middle school for reading (2 months) and maths (0.7 months).[5] Kogan and Lavertu (2021) analyse reading attainment in a statewide assessment in Ohio, and find that economic attainment gaps grew by 0.05 ES units from Autumn 2019 to Autumn 2020.[6] As the study authors note, however, this is likely an underestimate given the high levels of measurement error in the pupil-level disadvantage indicator, whereby some schools/districts are known to report all students as disadvantaged.

Engzell *et al.* (2020) examine attainment gaps in the Netherlands between students whose parents have differing levels of education.[7] This study examines attainment from national assessments immediately before covid-related closures (January/February 2020) and soon after schools returned (May/June 2020). Focusing on a composite

[2] Juniper Education (2020) also provide direct evidence of changes in the disadvantage gap using teacher assessments for Years 1 to 6. A clear strength of the Juniper study is that the data covers three periods: Autumn 2019, Summer 2020, and Autumn 2020. The disadvantage gap is defined in terms of the percentage of students who were assessed by their teacher to be achieving at expectation. The gap grew from Autumn 2019 to Summer 2020 by around 4.8 percentile points in maths and reading. This widening was reduced by roughly half in the Autumn term.
[3] We thank Katie Blainey from RS Assessment for providing the underlying data to Figures 7 and 8 so that these numbers could be calculated.
[4] While they are unable to provide any direct evidence, the authors compare FSM gaps in their Autumn 2020 study sample to national estimates of the FSM from Key Stage 1 assessments in 2019, concluding that '[i]t seems that the disadvantage gap is wider than earlier estimates' (NFER, 2021, p. 12).
[5] Elementary/middle schools in the data cover Grades 4–8. Note that Pier *et al.* (2021) report a substantial *narrowing* of gaps in high school—although this is beyond the scope of our paper. The outcomes are MAP and STAR assessments.
[6] As the study authors note, this is likely to be an underestimate, given the high levels of measurement error in the pupil-level disadvantage indicator, whereby some schools/districts are known to report all students as disadvantaged.
[7] One limitation of this high-quality study is the measure of disadvantage. Household are divided into three categories in terms of parental education. 92% of households fall into the 'high' category, which makes this quite a restrictive measure.

measure of attainment in Years 4 to 7, the paper finds an increase in gaps of around 1.2 percentile points (roughly 0.03 ES units, ± 0.007).[8]

Finally, a broader set of studies provides indirect evidence about the changes in economic attainment gaps, by comparing how covid-19 has affected average levels of attainment in schools (or neighbourhoods) that have high or low levels of disadvantaged pupils. In the UK, a study by the Department for Education examined attainment in reading and maths in Years 3–7 (DfE, 2021). Rather than estimating attainment gaps directly, the study models counterfactual outcomes to calculate 'learning loss' due to covid-19 and then compares levels of lost learning in schools with differing percentages of FSM pupils ('high', 'medium', and 'low').[9] The study finds that students in 'high' FSM schools had 0.3 months more learning loss than those in 'medium' and 'low' FSM schools. In the US, Renaissance (2020) found that schools with higher levels of disadvantaged pupils had greater levels of 'learning loss' compared to other schools (2 percentile points greater loss in maths, 1 percentile point in reading).[10] Similarly, Curriculum Associates (2020) compare estimates of average learning loss in schools from low or higher-income neighbourhoods. Across Grades 1 to 8 they find a small widening of economic attainment gaps for maths, but a small narrowing of gaps for reading.[11] Finally, Maldonado and De Witte (2020) report that schools in Belgium with higher percentages of students receiving financial support had greater learning losses in maths and Dutch.[12]

To summarise, evidence suggests that covid-19 led to widening income-based inequalities across a range of educational inputs. These inequalities seem to be widening attainment gaps. Schools and neighbourhoods with higher percentages of economically disadvantaged students have tended to be disproportionately affected by covid-19 in terms of average estimated learning loss. At the pupil level, direct estimates of gaps tend to show that they have widened for primary school students—although these estimates have some methodological limitations. The most relevant evidence for this study (DfE, 2021; RS Assessment, 2020) suggests that gaps may have grown by roughly 0.5 to 1 month's progress between Autumn 2020 and November/December 2020.

2. Documenting school responses and their impacts

Our second research focus is to document how schools and teachers responded to the huge disruptions caused by covid-19. Several studies surveyed teachers, asking them about their experiences of lockdown and the unexpected closure of schools. Of particular relevance for our study is evidence about how variation in school-level practices differed in terms of the economic circumstances of their pupils. For example, Cullinane and Montacute (2020) reported early in the initial lockdown that pupils from middle-class homes were almost twice as likely to be taking part in live or recorded lessons, compared to working-class peers (30% compared to 16%, see also Allen *et al.*, 2020, for evidence from later in the first lockdown). Similarly, schools with more affluent students were more likely to report using an online platform to help set and receive work (Cullinane and Montacute, 2020).

In May, NFER sent a survey to all primary and secondary schools in England. Around 9% of primary schools responded (Lucas *et al.*, 2020a).[13] The survey suggested that schools with higher levels of disadvantaged children had less regular contact with pupils: teachers in the most deprived schools were in regular contact with 50% of their pupils, compared to 67% for the least disadvantaged schools. Similarly, senior leaders in the most disadvantaged schools were more likely to report using workbooks or worksheets, and significantly less likely to be using live or pre-recorded video lessons.

Our second contribution is to build on existing evidence describing the variation in school responses during the initial lockdown (March–May 2020). We also extend existing literature by providing descriptions of how school responses varied when schools re-opened to all students in the Autumn term (September–December 2020).

---

[8] This assumes a normal distribution.

[9] 'High' means more than 25% of pupils are eligible for free school meals, while 'low' means less than 10% of pupils are eligible for free school meals.

[10] The comparison was between Title 1 schools, who have above-average levels of economically disadvantaged students, and others. This study examined Grades 1–8.

[11] The comparison is between Autumn 2020, and an average of three previous years (Autumn 2017, 2018, and 2019). The outcome is 'percentage of students who are two or more grades below expectation' on the i-ready diagnostic test.

[12] Other measures of disadvantage such as 'percentage of students who come from low SES neighbourhoods' and 'percentage of students with mothers who have low levels of education' found mixed results.

[13] Results were reweighted so that the survey sample reflected the population in terms of observed characteristics.

3. Linking school responses to changes in educational inequalities

Part 3 of our study examines associations between school responses to covid disruptions (both during the initial lockdown, and when schools re-opened) and economic gaps in attainment. As far as we are aware, no existing research has explored this relationship directly. Lucas *et al.* (2020b) use teacher surveys to report associations between schools' use of a virtual learning environment (VLE) and maintaining a broad curriculum with higher pupil engagement. Andrew *et al.* (2020), using parent surveys of time use, report a positive association between the use of 'active school resources' (online classes, video conferencing, and online chat) and more time spent on learning activities, after controlling for a range of family socioeconomic factors and home resources. While these studies provide useful evidence, they rely on behavioural proxies for learning (engagement and time spent). In addition to being proxies, these measures are often based on teacher or parent reports, which may be confounded by the fact that the people who are reporting on the behaviours are often responsible for bringing these behaviours about. This makes it difficult to interpret these analyses as clear evidence of a relationship between school responses and pupil learning. Our third contribution is therefore to provide direct, novel estimates of the association between school practice and inequalities in attainment, during covid-19.

Research questions

We have three research aims:

1. Quantify changes in inequalities in educational attainment since the onset of covid-19;
2. Describe how teachers and schools responded to the challenges of covid-19;
3. Explore associations between school responses and changes in attainment gaps.

These aims correspond to the following sets of research questions.

1. Quantify changes in inequalities since the onset of covid-19[14]

- RQ1a: How did the attainment gap between disadvantaged pupils and their peers change during the period of (partial) closures due to covid (November/December 2019 to September 2020)?
- RQ1b: How did the attainment gap between disadvantaged pupils and their peers change after a term of (largely) in-person schooling (September 2020 to November/December 2020)?
- RQ1(c): How has the attainment gap between disadvantaged pupils and their peers changed overall since the onset of covid (November/December 2019 to November/December 2020)?

2. Describe how schools responded to covid-related disruptions and their consequences

- RQ2a: How did schools respond during the initial closure period (March–May 2020) to support pupils working from home?
- RQ2b: How did schools respond during the re-opening period (September–December 2020) to compensate for lost instruction time and deal with ongoing covid-related disruptions?

3. Explore associations between disadvantage attainment gaps and school responses

- RQ3a: What was the association between school responses during the initial closure period (March–May 2020) and changes in the disadvantage attainment gap when children returned to in-person schooling (from November/December 2019 to September 2020)?
- RQ3b: What was the association between school responses during the re-opening period (September 2020–December 2020) and changes in the disadvantage attainment gap when children returned to in-person schooling (September–November/December 2020)?

The study plan is available on the EEF website at:
https://educationendowmentfoundation.org.uk/public/files/Projects/EEF_Covid_Partial_Closure_Impact_Response_3.1.pdf

---

[14] The study plan contains two additional questions, not addressed here, focusing on attainment in Spring 2021 and Summer 2021. If these assessments go ahead, we will provide addenda as per the study plan.

Ethics

FFT was responsible for school recruitment and data collection. Schools opted into the project via and online form that was completed by the headteacher. The online form was sent in an invitation email to 800 FFT Aspire schools that we believed would also have RS Assessment test data.[15]The email outlined the project aims and requirements and invited schools to complete an online form with the contact information required for the project. Additionally, it required headteachers signing up to tick a box confirming their school's participation in the project as well as agreeing the data already shared with FFT can be used for this research.

Due to the speed required for project setup, participating schools were required to be FFT Aspire customers, as the project relies on the data schools already share with FFT, and the terms and conditions schools sign up to for FFT Aspire to enable the research and analysis to take place. These terms and conditions enable FFT to use schools' data to undertake research.

The welcome email, along with the online form headteachers completed to be a part of the research, can be seen in Appendix A.

Data protection

The legal basis for processing the personal data used in this research project is legitimate interest Article 6(1).

FFT Aspire customers agree to Terms and Conditions and Terms of Use that outline how data shared with FFT can be used to undertake research into how education systems function. This is:

- Necessary for our customers' legitimate interests (analysis of pupil performance and requirements, school performance, staff performance, and ensuring equality of opportunity and treatment of pupils);
- Necessary for our legitimate interests (to be able to undertake research for public benefit)
  Link to privacy notice can be found here: **https://fft.org.uk/privacy/**.

Participating schools have volunteered to take part in this research and understand that the data shared with FFT Aspire will also be used for this research.

The personal data for this research will only be accessible by FFT colleagues through FFT's IT infrastructure. FFT is certified and accredited to ISO 27001 (the international standard for Information Security) and is registered as a data controller with the Information Commissioner's office. FFT is also certified to the government's Cyber Essentials Plus standard and completes an annual IT Health Check of our complete internal- and external-facing IT systems, websites, and infrastructure which is reviewed and approved by the Department as part of our accreditation to receive DFE ASP data.

Data being used for this project is defined as Customer Data (as noted in FFT's privacy policy). This means that the customer is the Data Controller and FFT is the Data Processor. The data is only accessible by FFT colleagues through FFT's IT infrastructure and all pupil data is encrypted both in transit and at rest. No pupil data was shared with collaborators at the EEF or Teacher Tapp.

Customer data is provided by schools for use within the Aspire platform, subject to the terms and conditions agreed between the customer and FFT. Data retention periods for Customer data are detailed in the FFT privacy notice. Customer Data will be deleted no later than 48 months after termination of a customer's subscription term.

The data for this project will not be available for archiving.

[15] FFT provided RS Assessment with a list of schools that are FFT Aspire customers. RS Assessment then used this list to indicate schools that had purchased RS Assessment tests in previous years and in 2020.

Project team

FFT were responsible for the recruitment of schools, collecting pupil assessment data, and teacher survey data. This was largely led by Laura James. Pupil data collection permission was followed up by Ruth Jameson as part of FFT's School Support Team. Dave Thomson, Natasha Plaister, and Dave Bibby from FFT were responsible for analysing data. Ben Weidmann from the EEF wrote the report with assistance from Rob Coe and FFT colleagues. Becky Allen designed and analysed the teacher surveys. The project was supported by EEF colleagues Jennifer Stevenson, Celeste Cheung, and Jamila Boughelaf.

Methods

Study design

Table 2: Study design

Design		Observational study
Primary unit of analysis		School
Primary outcome 1	Variables	Maths and reading attainment in Years 2–6, in September/October 2020 ('missed Summer assessment')
	Measures (instruments, scale, source)	PIRA, PUMA, NTS (RS Assessment tests; standardised scale*)
Primary outcome 2	Variables	Maths and reading attainment, in Years 2–6, in November/December 2020 ('Autumn 2020 assessment')
	Measures (instruments, scale, source)	PIRA, PUMA, NTS (RS Assessment tests; standardised scale*)
Baseline for primary outcomes	Variables	Maths and reading attainment, using RS Assessment tests in Years 1–5, in Autumn 2019
	Measures (instruments, scale, source)	PIRA, PUMA, NTS (RS Assessment tests; standardised scale*)

*The standardisation is described in the Measures section.

Timeline

Dates	Activity
July 2020	School recruitment and first teacher survey administered
September/October 2020	Schools administer Spring 2020 RS Assessment tests
November/December 2020	Schools administer Autumn 2020 RS Assessment tests
December 2020	Second teacher survey administered
January 2021	Collection of assessment data from schools
March 2021	Draft report delivered
March/April 2021	Third teacher survey administered and schools receive school reports
June 2021	Schools administer Summer 2021 assessments
August 2021	Addendum to report

Participants, data, and sample sizes

This section begins by providing an overview of the different types of data we use in the report. We then describe our recruitment strategy and the analysis samples we use to address each research question.

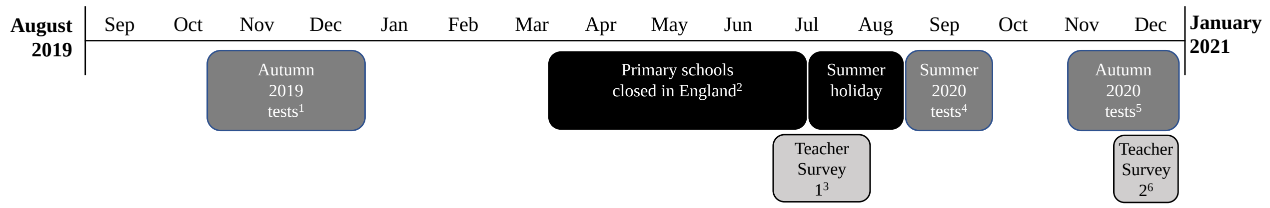
Overview of data sources and timeline

This study draws on multiple data sources from several time points. We collected attainment data from three time periods:

- November/December 2019 (**'Autumn 2019'**);
- September 2020 (labelled **'Summer 2020'** as these were tests that were originally scheduled for Summer. Because of covid-19 they were administered when students returned to in-person schooling in September 2020);
- November/December 2020 (**'Autumn 2020'**).

Information on school responses come from two separate surveys: Teacher survey 1 (in July 2020) and teacher survey 2 (in December 2020). *Figure 1* provides an overview of our data sources.

Figure 1: Overview of data sources



Notes: [1] Autumn 2019 tests were primarily conducted in November (42%) and December (52%). The remainder were almost all in October 2019.
[2] Schools were closed on March 25th. During the period from March 25th to June 1st, most schools were only open to a very small group of children. From June 1st 2020, many primary schools re-opened for children in Reception, Year 1, and Year 6.
[3] Teacher survey 1 was focused on the initial period of school closures (March–May 2020) and was administered in July 2020.
[4] Summer 2020 tests were conducted in-person, when schools re-opened in September for the start of the 2020–21 school year.
[5] Autumn 2020 tests were almost all administered in December (65%) or November (32%).
[6] Teacher survey 2 was focused on how schools responded to the return of face-to-face schooling in the Autumn term of 2020–21. It was administered in December 2020.

Recruitment

Recruitment of schools

This study focuses on state-funded primary schools in England. School recruitment ran from May 2020 to July 2020. There were two stages of school selection. First, we used a list provided by RS Assessment to identify schools that had administered RS Assessment tests in reading and/or maths, in Year Groups 1, 2, 3, 4, or 5 in Autumn 2019.[16] Second, we contacted these schools and asked if they would be willing to:

1. Administer tests in September 2020 and November/December 2020;
2. Allow us to collect test data from RS Assessment;
3. Allow us to collect pupil context data from their school management information system (MIS).[17]

145 schools volunteered and were offered a financial incentive to participate in the study,[18] along with a tailored report comparing attainment changes in their school with the broader research sample. Eight schools withdrew before the 2020–2021 academic year[19] and five schools failed to provide data. In total, 132 schools supplied

[16] Our original study plan stated that we would be examining Years 2–5 (in Autumn 2019). We originally did not include Year 1 as we thought that the sample size of schools administering assessments in Year 1 would be much smaller than other year groups. However, when we received the data from schools it became clear that we had as much data for Year 1 as we had for Year 2. We therefore decided to include Year 1, in an effort to provide a fuller picture of achievement across primary school.
[17] Recruitment materials are presented in Appendix A.
[18] £250 to be paid in January 2020, and a further £100, paid at the conclusion of the study in May/June 2021
[19] The reasons for drop out were due to schools not planning to use RS Assessment tests and a misunderstanding from the original communication.

some data for the project. The samples we use to answer specific research questions are described later in this section.

Recruitment of teachers

We recruited a sample of teachers from within the set of 145 volunteer schools. In July 2020 we contacted volunteer schools and asked them provide email addresses for up to three teachers in each relevant year group (Years 1 to 5 in 2019–2020). We received 950 email addresses.

Shortly after assembling the list of teacher email addresses, we disseminated Teacher Survey 1 using a Google form. The questions, which are presented in Appendix C, focused on the details of the teacher's remote learning provision during March–May 2020. In total, 539 teachers from 137 primary schools responded to the survey.

The same set of teachers were sent Teacher Survey 2 at the end of December, 2020. The questions for Teacher Survey 2, which are presented in Appendix D, focused on how schools responded to the challenges of returning to face-to-face instruction, while covid-19 continued to cause disruptions. In total, 454 teachers from 93 schools responded to the survey.

Table 2 illustrates the number of responses per school for each survey. In some schools, we were able to get a relatively large number of responses: for example, in 17 schools we had eight or more teachers respond to Teacher Survey 1. Our approach here was to maximise information, i.e. to get responses from as many teachers as possible in our sample of recruited schools (while trying to complete recruitment and surveying as quickly as possible).

Table 2: Number of teacher survey responses per school

Number of teacher responses per school	Survey 1: count of schools	Survey 2: count of schools
1	39	16
2	19	10
3	22	11
4	9	18
5	15	9
6	12	7
7	4	6
8 or more	17	16
Total	**137**	**93**

Overview of analysis samples

Analysis sample for Research Question1

To address RQ1, the composition of our analysis sample ideally needed to be constant at each of the three assessment points. If the composition of the sample were allowed to change over time, then any changes in estimates of the disadvantage attainment gap could be due to sample compositional changes, rather than shifts in the underlying attainment gap between FSM6 pupils and their peers.

Our starting sample was defined by pupils who:

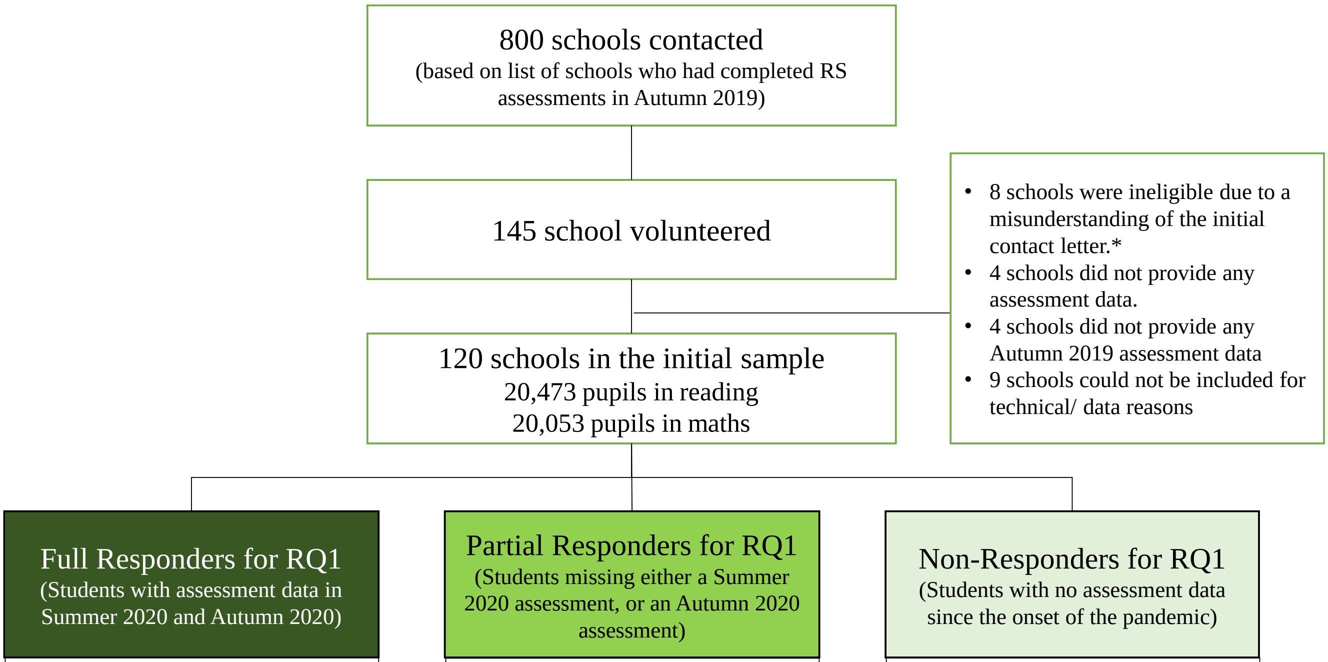
- had a valid assessment in reading or maths in Autumn 2019; and
- were on roll during the Autumn term 2020 and had a disadvantage (FSM6) flag (yes/no) present.

From the initial set of 145 schools, there were a total of 20,473 students who were part of the starting sample for reading, and 20,053 for maths. In Autumn 2020, many pupils faced disruptions, including to their reading and maths assessments. This is discussed in the section on missing data. Here we note that, for the purposes of RQ1, we specified three categories of pupil:

- Full responders: those for whom both a Summer 2020 and an Autumn 2020 result is observed;
- Partial responders: those for whom either a Summer 2020 or an Autumn 2020 result is observed (but not both);
- Non-responders: those for whom no test results are (yet) observed since the onset of covid-19.

These samples are summarised in Figure 2.

Figure 2: Overview of sample for RQ1



Full Responders for RQ1

	Reading	Maths
Yr2	1,295	1,295
Yr3	1,217	1,261
Yr4	1,911	1,950
Yr5	2,190	2,273
Yr6	1,746	1,755
Total	8,359	8,534

Partial Responders for RQ1

	Reading	Maths
Yr2	1,474	1,558
Yr3	1,563	1,555
Yr4	2,109	1,770
Yr5	1,736	1,442
Yr6	1,659	1,662
Total	8,541	7,987

Non-Responders for RQ1

	Reading	Maths
Yr2	444	417
Yr3	434	567
Yr4	632	702
Yr5	664	663
Yr6	1,399	1,183
Total	3,573	3,532

Notes: *Misunderstandings were related to (i) not having PIRA/PUMA/NTS data from Autumn 2019 or (ii) schools who signed up, but who were intending to switch to PIXL assessments.

Analysis sample for Research Question 2

We used all available responses from Teacher Surveys 1 and 2 to address RQ2. Starting from a possible sample of 145 schools, we received data from 137 schools in Teacher Survey 1 (n_{teacher} = 539) and 93 schools in Teacher survey 2 ($n_{\text{teacher}} = 454$).

Analysis sample for Research Question 3

Addressing Research Question 3 required overlap between assessment data and teacher survey data. Specifically, for RQ3a we required overlapping data from Teacher Survey 1 and assessment data from both Autumn 2019 and Summer 2020. The resulting overlap is presented in Figure 3. For RQ3b, we required overlap between Teacher Survey 2 and assessment data in both Summer 2020 and Autumn 2020. The overlap is presented in Figure 4.

Figure 3: Overview of sample for RQ3a

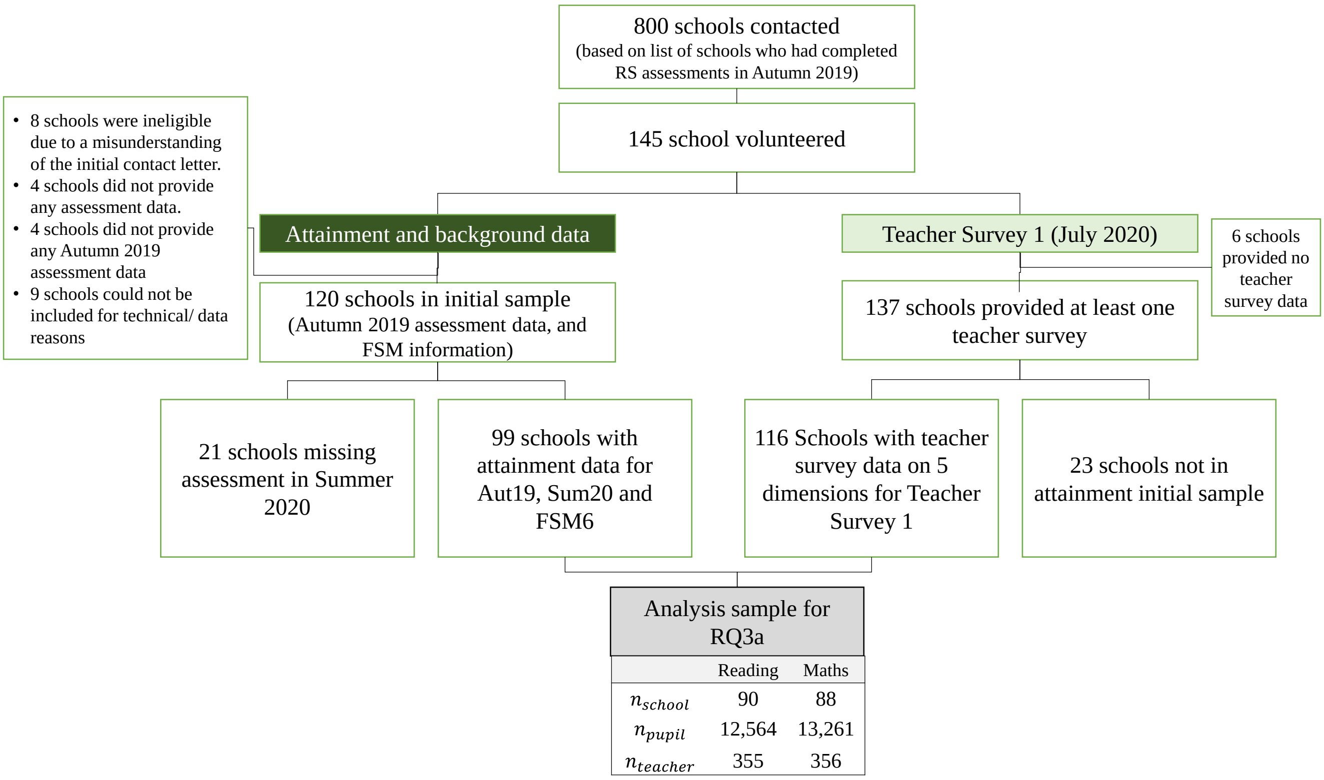
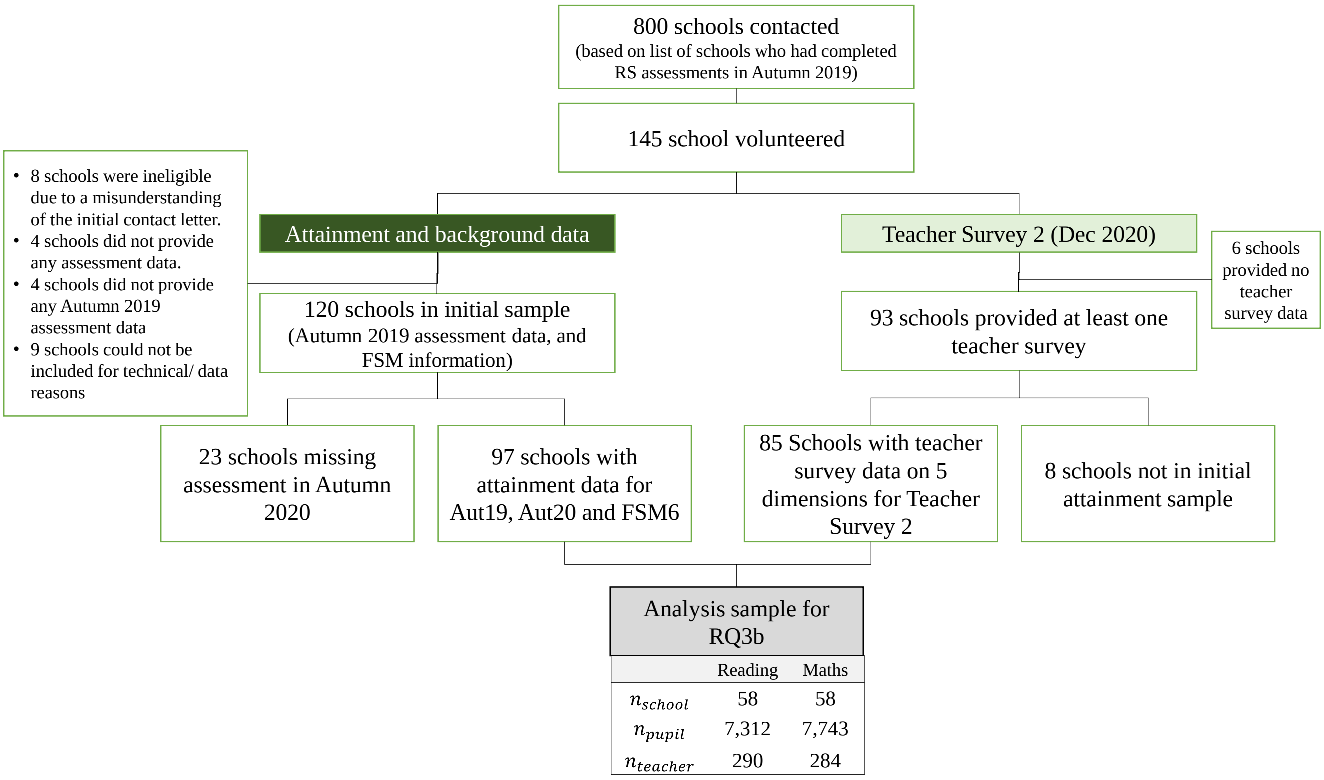


Figure 4: Overview of sample for RQ3b



Comparing analysis samples to broader samples/populations

Comparing our research samples to all state-funded primary schools

We begin by comparing observed school characteristics of our analysis samples, to the population of state-funded primary schools in England. This information is summarised in Table 3.

The schools in our RQ1 sample, which we use to estimate disadvantage gaps, are similar in terms of the percentage of disadvantaged students (34% compared to 30%) and extremely similar in terms of mean prior attainment (mean of 16.0 compared to 16.3 points). Our RQ1 sample under-represents schools from the north of England (22% compared to 32% in the population) and from London (8% compared to 11%). The sample over-represents students whose first language is something other than English (29% compared to 21%).

Table 3: Characteristics of schools in RQ1, RQ2, and RQ3 compared to the population of state-funded primary schools

	RQ1	RQ2a	RQ2b	RQ3a	RQ3b	**Population**
% junior schools	8%	10%	6%	8%	8%	**7%**
% London schools	8%	12%	10%	9%	6%	**11%**
% northern schools	22%	24%	27%	25%	23%	**32%**
Mean cohort size (pupils)	46.6	45.9	45.5	47.0	47.4	**41.7**
Mean prior attainment (Key Stage 1 average points)	16.0	16.2	16.0	16.1	15.9	**16.3**
% disadvantaged pupils	34%	33%	35%	34%	35%	**30%**
% first language other than English	29%	26%	31%	28%	32%	**21%**
% expected standard in reading, writing, and maths	62%	62%	62%	63%	62%	**66%**
% higher standard in reading, writing, and maths	9%	10%	10%	9%	9%	**11%**
Number of pupils (all year groups, thousands)	38	44	30	31	22	**4,604**
Number of schools	120	139	93	95	64	**15,580**

Notes: for a description of recruitment and participant flow, see Figure 2, Figure 3, and Figure 4. All the study samples in this table refer to the sample we use for analysis, for example, RQ1 = 'responder sample for RQ1'. The 'population' represents all state-funded primary schools and data comes from the UK government **website**.

Table 3 also compares the samples we use to address RQ2a and RQ2b with the population. These samples again somewhat under-represent schools from the north of England, especially our RQ2a sample (24%, compared to 32% in the population). However, in terms of disadvantage (as measured by FSM6) the sample is fairly similar to the population (34% compared to 30%), and the mean prior attainment is very similar.

Finally, Table 3 compares the observed characteristics in the population with the samples we use to address RQ3. As our RQ3 sample represents the overlap of RQ1 and RQ2, the patterns here are familiar. The samples under-represent schools in London and the north of England and has slightly above-average levels of disadvantage. Mean prior attainment is very close to the population average.

Overall, based on the characteristics for which we are able to make comparisons, our analysis samples are all quite similar to the population, notwithstanding some differences in geography.

Comparing respondents to our teacher surveys to a broader sample of teachers in England

Next, we turn to the question of whether our samples are unusual in terms of the ways schools responded to remote learning (Teacher Survey 1) and the return to in-person instruction (Teacher Survey 2).

Fortunately, several of our survey questions were also asked to a wider sample of teachers in England via the daily survey app called Teacher Tapp. In Table 4 we show that the school practices in our sample are very similar to those reported in the wider Teacher Tapp sample. Of 20 common questions, none had mean differences of greater than 15%. The question with the biggest difference was the percentage of teachers who reported spending no extra time on English and maths (38% in our sample, 53% in the broader Teacher Tapp sample). Overall, the similarity between the two samples gives us some confidence that our sample was not unusual in terms of their response to covid-related disruptions. More details on the Teacher Surveys are provided in the section on Measures, and Appendices C and D.

Table 4: School practices in study schools compared to a broader sample of teachers in England.

Question	Response categories	RQ2 sample	Teacher Tapp
Did you provide students with a timetable to follow each day?	No	33%	34%
	Yes, but it was only a suggested timetable and pupils were not required to follow it	52%	54%
	Yes, we asked pupils to follow a daily timetable	15%	12%
	Number of responses	539	1449
How frequently did you suggest that parents or pupils send in work?	I gave no suggestion about how frequently work should be sent in	29%	39%
	Once a week (or less)	21%	11%
	Several times a week	17%	18%
	Every day (or after each lesson)	33%	33%
	Number of responses	539	1436
Did children (i.e. not the parents) typically speak to you (i.e. their class teacher) on the phone?	No, children typically did not speak to their class teacher	34%	46%
	Yes, I spoke to them each once or twice during the half term	42%	30%
	Yes, I spoke them to each about once a week (or more frequently)	24%	23%
	Number of responses	539	1431
Did you try teaching a 'live' lesson to your class during the first half of the Summer term (with students either speaking or on mute)?	No	52%	61%
	No, but I did pre-record a video of myself for them to watch	35%	29%
	Yes (any number of live lessons)	12%	10%
	Number of responses	539	1442
Is your school making greater or less use of small group and individual face-to-face interventions (for example, run by TAs) so far this year?	There are more intervention groups running than in a normal year so far	25%	20%
	Similar number of intervention groups running compared to a normal year so far	33%	38%
	There are fewer intervention groups running compared to a normal year so far	41%	42%
	Number of responses	423	1611
How many extra minutes a day (over and above a normal year) are you spending on English and maths right now to cope with missed learning as a result of shutdown and absences?	Zero minutes—we have not reallocated lesson time at all	38%	53%
	Extra 15 minutes	23%	15%
	Extra 30 minutes on English and maths	25%	21%
	Extra 45 minutes (or more) on English and maths	14%	11%
	Number of responses	403	1432

Are the schools in the RQ3 analysis different from our RQ2 sample?

Finally, we examine whether the school response variables we analyse differ from RQ2 (in which we are able to use the full teacher samples) and RQ3 (where we require overlap between teacher survey data and attainment data).[20] These tables suggest that our analysis samples for RQ3 are very similar to the full teacher survey samples in RQ2, in terms of variables describing schools' responses to covid-19 disruptions.

Table 5: Comparing school responses from the full sample of Teacher Survey responses (RQ2), and analysis sample for RQ3

	Analysis sample for RQ3a	Analysis sample for RQ2a
School response variable (mean)		
Intensity of online platform use	2.24	2.41
Timetabling	0.73	0.82
Regularity of student work submissions	1.44	1.54
Speaking to students on the phone	0.87	0.90
Live or pre-recorded lessons	0.60	0.60
Number of schools*	90	137

Note:*count of schools in either the maths or reading analyses.

[20] See the Measures section for a description of variables.

Table 6: Comparing school responses from the full sample of teacher survey 2 responses, and analysis sample for RQ3b

	Analysis sample for RQ3b	Analysis sample for RQ2b
School response variable (mean)		
Absence provision for individuals (recording or live lessons)	0.25	0.24
Closure provision for classes (recording or live lessons)	0.89	0.97
Curriculum reductions	1.34	1.44
Extra time	0.25	0.28
Small group interventions	1.06	1.13
Number of schools*	58	93

Notes:*count of schools in either the maths or reading analyses

Measures

Attainment data

All schools in the sample used tests in reading and/or maths provided by RS Assessment.[21] Three tests were used: Progress in Reading Assessment (PIRA); Progress in Understanding Mathematics Assessment (PUMA); and NTS Assessments (national test style) reading and maths papers.

PIRA, PUMA, and NTS are widely used termly assessments taken by over 6,000 primary schools in England. In Autumn 2020, RS Assessment introduced a new suite of PIRA and PUMA tests known as PIRA 2021 and PUMA 2021.

Schools in the sample therefore had the choice of three types of standardised test in reading and three types of standardised test in maths. In reading the tests were:

i) PIRA 2016;
ii) PIRA 2021;
iii) NTS Reading.

In maths the tests were:

i) PUMA 2016;
ii) PUMA 2021;
iii) NTS Maths.

The tests take around 45–50 minutes and provide coverage of the revised national curriculum for each year group. The tests can be taken online, although we believe the majority were taken using pen and paper. Similar to the arrangements for Key Stage 2 tests, all pupils in a year group are typically expected to take the test, with the exception of any pupils who were absent or were unable to access the test for another reason. Scripts were marked by teachers using published mark schemes and the scores for each pupil were entered into MARK, an online reporting and analysis tool provided by RS Assessment, or Aspire Pupil Tracking, a reporting tool provided by FFT.

Standardising attainment data

Although all RS Assessment tests report standardised scores, the tests were standardised at different times. We considered three approaches to equating scores from different tests (PIRA, PUMA, and NTS) conducted at different periods (Autumn 19, Summer 20, Autumn 20):[22]

1. Standardise scores. In accordance with the study plan this is our preferred approach and is described below.

[21] **https://www.risingstars-uk.com/rs-assessment**
[22] See Figure 1 for detailed information on dates.

2. Calculating ranks, within each combination of:
    a. test (for example, PIRA16);
    b. time period (for example, Autumn19);
    c. year group (for example, Year 2).
3. Calculating ranks, within each combination of: test-time–year group *and* school.

Note that the impact of equating is discussed in the Robustness section. We find that our results are largely insensitive to our method of standardising scores.

Standardised scores (main approach)

For all tests (PIRA, PUMA, and NTS) we received scores from RS Assessment that had been normed relative to a representative sample of students in England.[23] The technical manuals report reliability coefficients (Cronbach's Alpha) around 0.9 for all the tests we analyse (McCarty and Cooke, 2015; McCarty and Ruttle, 2016).[24]

Let Y_{ijgta}^{Norm} be the scores we received from RS Assessment for pupil i, in school j, for year group g (defined in terms of their year group in 2020–2021), at time t (Autumn 2019, Summer 2020, Autumn 2020), measured by assessment a (PIRA16, PIRA21, PUMA16, PUMA21, NTS Reading, NTS Maths). Because the norming was done in different years, we take the extra step of adjusting scores to have the same means and standard deviations of the large samples of students taking these tests in the three assessment points we analyse. Specifically, we calculate:

$$Y_{ijgta} = \frac{Y_{ijgta}^{\text{Norm}} - \hat{\mu}_{tga}}{\hat{\sigma}_{tga}}$$

where $\hat{\mu}_{tga}$ is the mean score on assessment a at time t in year group g in the full sample of children who sit RS Assessment tests. The number of children in these broader samples range from 11,727 to 19,794 for each combination of assessment-time–year group.[25] This leaves us with our main measure of attainment Y_{ijgt}^{k}, in which the 'assessment' subscript a is replaced by a 'subject' superscript k which indicates reading or maths.

Robustness checks: ranks

To check whether our results are robust to different methods of scaling, we converted attainment scores into ranks. We did this in two different ways. First, we looked at ranks within each combination of assessment-time–year group. Let n_{ijgta}^{k} be the ordinal rank of pupil i in year group g, at time period t, in assessment a (for subject k) in terms of Y_{ijgta}^{k}. We calculate:

$$R_{ijgt}^{k} = 1 + \text{floor}\left(100 \times \frac{n_{ijgta}^{k} - 0.5}{N_{gta}^{k}}\right) \quad \text{(A)}$$

where N_{gta}^{k} is the number of students in our analysis sample at time t for year group g in assessment a. In order to make the results easily comparable across scaling approaches we converted the rank measure R_{ijgt}^{k} into standard deviation units, using the normal CDF (Φ):[26]

[23] Note that this norming process took place in different years. For example, PIRA16 norming took place in 2013–2014; PUMA16 norming took place in 2014–2015. PIRA21 and PUMA21 were both normed in 2020.

[24] During the standardisation process, the PIRA tests were correlated with teacher assessments. Correlations were in the range 0.72 to 0.79. In PUMA, the correlation between the Summer Year 6 test and the national Key Stage 2 maths test was 0.83.

[25] We were unable to source $\hat{\mu}_{tga}$ and $\hat{\sigma}_{tga}$ for $t = \text{Summer 2020}$ and $a = (\text{NTS Reading, NTS Maths})$. In both these cases, we used estimates of the mean and standard deviation of Y^{Norm} from our own sample.

[26] As noted in the Robustness section, we also ran our analyses using ranks R. This made no difference to the pattern of results.

$$Y_{ijgt}^{k,\text{Rank}} = \Phi\left(\frac{R_{ijgt}^{k}}{100}\right) \text{ (B)}$$

As a second robustness check, we calculate ‘within-school ranks’. We follow the same procedure outlined in equations (A) and (B), but calculate ranks within school:

$$R'^{k}_{ijgt} = 1 + \text{floor}\left(100 \times \frac{n_{ijgta}^{k} - 0.5}{N_{jgta}^{k}}\right) \ (X3)$$

where N_{jgta}^{k} is the number of students in school j who provided data for subject k at time t, and n_{ijgta}^{k} is the rank of pupil i within their year group g at school j.

School responses

This section describes Teacher Surveys 1 and 2. The topics of the surveys were:

- Survey 1: remote learning provision from March to May 2020;
- Survey 2: strategies used in the Autumn 2020 term to deal with: (i) disruptions and absences due to covid; (ii) lost learning during the initial lockdown.

Teacher Survey 1

Questions about the initial lockdown were based on the idea that several factors may influence the effectiveness of remote education, and that schools may make different choices in terms of these factors. Some choices were appreciably more expensive or difficult than others, so it seemed valuable to know whether there were associations between choices and pupil outcomes. Many of the questions we asked were tested and refined on a wider primary teacher sample using the Teacher Tapp survey platform. We asked a total of 16 multiple-choice questions, summarised below. The full questions, along with the response options are in Appendix C.

Overview of Teacher Survey 1 Questions:

1. How was work set? (for example, website download, online learning platform, through live video lesson)
2. How did pupils send completed work back to you? (for example, email, online platform, Facebook)
3. How far was the school's pre-existing curriculum reduced, paused, or stopped?
4. Which of the following resources were used as part of the home learning you set for students? (for example, subscription site, free online resources, Oak National Academy)
5. To what extent were literacy or English resources adapted or created for parents to use?
6. To what extent could literacy or English resources be completed by children without supervision from their parents?
7. Did you have any form of daily registration for students learning at home?
8. How much of a structured daily timetable did you use? (for example, required, suggested, not used)
9. During the first half of the Summer term, could pupils: take part in live lessons (with or without talking); take part in live or chat-based social interaction; watch a pre-recorded video of their teacher; none of these?
10. How often did children (i.e. not the parents) typically speak to you (their class teacher) on the phone?
11. Did you share examples of student work somewhere for all students to see? (for example, in newsletter, on website, email, social media, or not)
12. What kind of feedback did you give on pieces of work submitted by pupils? (for example, general praise/encouragement, specific feedback to support learning on some/all work, or none given)
13. How frequently did you suggest that parents or pupils send in work?
14. How much do you agree with the following statement: ‘During lockdown whilst most of my students were learning at home, it was easy for me to monitor who was, and wasn’t, completing work.’
15. How much do you agree with the following statement: ‘When setting work for remote learning, I found it difficult to differentiate to the lowest attainers in my class.’
16. Overall, how good do you feel the home learning experience was for your class?

As many of the questions were inter-related, we began our analysis by examining how responses co-varied. This involved two steps. First, we recoded the survey responses from the 16 questions into 29 binary and ordinal variables (as shown in *Table 15* in Appendix C). We then used exploratory factor analysis to condense these 29 variables into a smaller number of underlying latent variables that describe primary school practice during the March–May lockdown. This factor analysis was performed at the level of the individual teacher response, using a principal factors method (with rotation) to reduce our 29 variables to just eight factors. *Table 16* in Appendix C shows the rotated factor loadings, which gives an indication of which variables tend to co-vary.

The factor analysis did not produce a particularly clear factor structure, nor were the factors obviously interpretable. Our goal was to derive measures that captured important and interpretable aspects of school practice—which is partly a matter of judgement. Also relevant to this was the within-school intra-cluster correlation (ICC), an indicator of the extent to which teachers in the same school give similar responses to a question, relative to the variation between different schools.

For these reasons, we used the factor analysis to create a small set of interpretable, ad hoc variables to describe a school's provision for remote learning, focusing on those factors that seem to be consistent across a school.[27] Table 7 describes the five variables we use in our analysis.

Table 7: School Response Variable for Teacher Survey 1

Response variable	What it measures	Scale	Question*	ICC	Agreement°
Phoning students	Did teachers speak to students (not parents) directly on the phone? 1 = either one or twice during the half term; 2 = about once a week or more.	0–2	Q10	0.465	58%
Platform use	Did schools use a platform: to set work; to receive work; to share examples of student work; for daily registration; for text chat interactions; for live audio/video chat interactions? One point for each.	0–6	Q1, Q2, Q4, Q7, Q11	0.531	67%
Timetable	What timetabling approach did schools have? 0 = no timetable; 1 = a suggested timetable; 2 = instructed daily timetable.	0–2	Q8	0.337	54%
Videos/live lessons	Did teachers have pre-recorded video lessons (scored 1) or live lessons (scored 2)? If neither, the variable is coded as zero.	0–2	Q9	0.267	56%
Work submission	How often did teachers expect work to be submitted? 0 = no recommendation about work submission; 1 = suggested work submission once a week or fortnight; 2 = several times a week; 3 = every day.	0–3	Q13	0.410	47%

Notes: *See Appendix C for full list of questions. °This measures the percentage of times that pairs of teachers *within the same school* submitted the same set of responses on the questions underlying each of the five variables. This was based on 3136 pairs of responses from 98 schools. In the case of platform use (which is on a scale of 0 to 6) we coded teachers as submitting the same response if they were within one point of each other.

Teacher Survey 2

Survey 2 asked teachers about the way schools responded in the Autumn 2020 term, a period in which covid-19 cases were still prevalent. Given the level of disruption among schools and the demands placed on teachers, we decided to create a shorter questionnaire to minimise the burden on teachers (seven questions, one of which had multiple parts). These questions are listed in Appendix D. With fewer questions, we judged that it was not necessary to conduct a factor analysis to explore how to reduce the number of variables. This was a deviation from our study plan, driven by the changing circumstances. Instead, we simply combined questions that focused on the same element of school responses, creating a composite measure. For example, we combined all three prompts that asked about whether schools had increased their use of small group interventions, in an effort to

[27] The goal of providing *ad hoc* interpretable measures was specified in our study plan, available **here**.

compensate for lost learning (see Appendix D for details). The correlation of items at the school level is presented in *Table 18*, in Appendix D.

We ultimately created five *ad hoc* variables for analysis, summarised in Table 8. The table includes intra-class correlation (ICC), which were used to check whether practices were reasonably consistent within schools.[28]

Aggregating teacher responses into school-level variables

For each school, we calculate the mean of the teacher responses. It is worth noting that within schools there were disagreements between teachers as to what 'school practice' looked like. Disagreements within a school could be due to measurement error or reflect genuine differences across teachers and classes. In both cases, these disagreements act as a barrier to our analysis of how these practices are associated with learning and changes in attainment. We summarise these disagreements in two ways, using the intracluster correlation (ICC) and a measure of 'agreement' (see Tables 7 and 8). In both cases, we find that there is some variation in teacher reports within schools.

[28] Due to very low ICC, we dropped a survey question asking teachers 'In which of these subjects have you returned to (re)teach topics and skills that the class had missed as a result of lockdown during Summer term 2020?'.

Statistical analyses

This section describes our statistical methods. All analyses are run in R.

Research Question 1

Let Y_{ijgt}^{k} be a standardised attainment score for pupil i, in school j, for year group g (defined in terms of their year group in 2020–2021), at time t on subject k (reading or maths). The standardisation of attainment scores is discussed in the Measures section. Let F_i be a binary indicator equal to 1 if student i was classified as FSM6 in October 2020.[29]

We define the following three time points (see *Figure 1* for an overview of the timeline):

- T_0 = December 2019 (modal date of Autumn 2019/20 assessment);
- T_1 = September 2020 (modal date of Summer 2020 assessment);
- T_2 = December 2020 (modal date of Autumn 2020/21 assessment).

Our focus is the attainment gap between FSM6 pupils and their peers. We define this as G_{tg}^{k}: the difference in Y_{ijgt}^{k} between the mean attainment of FSM6 pupils and their peers for year group g:

$$\mathrm{G}_{tg}^{\mathrm{k}} = \frac{\Sigma_i Y_{ijgt}^{k} \cdot (1 - F_i)}{\Sigma_i (1 - F_i)} - \frac{\Sigma_i Y_{ijgt}^{k} \cdot F_i}{\Sigma_i F_i} \quad (1)$$

Our first research aim is to estimate G_t^k (the average attainment gap in primary schools for subject k) at different points in time, and then to estimate this quantity for different year levels, i.e. to estimate G_{tg}^{k}.

Focusing on the year-level specific estimates, let the change in G_{tg}^{k} from December 2019 to September 2020 be given by:[30]

$$\Delta_{T_1 g}^{\mathrm{k}} = G_{T_1 g}^{k} - G_{T_0 g}^{k} \ (2)$$

We estimate these quantities using statistical models. $\Delta_{T_1 g}^{\mathrm{k}}$ can be estimated using model 1a. We fit this model once for each year level (and subject k) using the lme4 package. In other words, we fit the following model 10 times:

$$Y_{ijT_1}^{k} - Y_{ijT_0}^{k} = \alpha_j + \delta_{T_1 g}^{k} + \Delta_{T_1 g}^{\mathrm{k}} F_i + \mathrm{month}_{iT_0} + \beta \mathrm{month\_difference}_i + e_{ijg}^{k} \quad \text{(model 1a)}$$

$$\alpha_j \sim N(0, \sigma_\alpha^2)$$

$$e_{ijg}^{k} \sim N(0, \sigma^2)$$

In each case:

- $\delta_{T_1 g}^{k}$ is the average change in Y (from to T_0 to T_1) for non-FSM6 pupils in year level g in subject k;
- $\Delta_{T_1 g}^{\mathrm{k}}$ is the average change in Y (from to T_0 to T_1) for FSM6 pupils in year level g in subject k.

[29] We expect the proportion of children who are designated FSM6 in October 2020 to be larger than the proportion in Autumn 2019, due to covid-19. This change in the composition of the two groups (FSM6 and non-FSM6) could influence our estimate of how the FSM gap has changed over time. To control for this, our headline results define FSM6 at a single point in time—October 2020. This ensures that the percentage of FSM6 children is constant at different time points and allows us to focus on how attainment in these two groups has changed. It may have been preferable to focus on FSM6 status in Autumn 2019, but unfortunately, as noted in the study plan, we did not have access to this data.

[30] In the study plan, the start date was listed as November 2019. However, the modal month of Autumn assessments in 2019–2020 turned out to be December.

- month_{iT_0} is a fixed effect for the month in which the Autumn 2019 assessment was taken by pupil i in subject k;[31]
- $\text{month\_difference}_i$ is the length of time, in months, between the administration of the Autumn 2019 and Summer 2020 assessments;
- α_j is a random effect for school j, to account for clustering of pupils within schools;
- e_{ijg}^k are normally distributed disturbances at the pupil level.

To present readers with a high-level summary of our results, we also estimate $\widehat{\Delta}_{T_1}^{\text{k}}$—an average change in disadvantage gaps across Years 2 to 6 for outcome k (from Autumn 2019 to Summer 2020). To see whether it was sensible to summarise the data in this way, we conducted likelihood ratio tests comparing a constrained model (1a') to an unconstrained model (in which $\Delta_{T_1g}^{\text{k}}$ can deviate by year level, i.e. model 1a). These tests are reported in Appendix E. We do not find evidence against the hypothesis that changes in disadvantage gaps are similar across year levels.[32]

$$Y_{ijT_1}^k - Y_{ijT_0}^k = \alpha_j + \delta_{T_1g}^k + \Delta_{T_1}^{\text{k}} F_i + \text{month}_{iT_0} + \beta\text{month\_difference}_i + e_{ijg}^k \quad (\text{model 1a}')$$

$$\alpha_j \sim N(0, \sigma_\alpha^2)$$

$$e_{ijg}^k \sim N(0, \sigma^2)$$

Our analysis for research question 1b is entirely analogous, but compares different time points, for example,

$$Y_{ijgT_2}^k - Y_{ijgT_1}^k = \alpha_j + \ \delta_{T_2g}^k + \Delta_{T_2g}^{\text{k}} F_i + \text{month}_{iT_1} + \beta\text{month\_difference}_i + e_{ijg}^k \quad (\text{model 1b})$$

$$\alpha_j \sim N(0, \sigma_\alpha^2)$$

$$e_{ijg}^k \sim N(0, \sigma^2)$$

Similarly, when comparing gaps across the entire period ($T_0 \rightarrow T_2$) we use the same model structures (which we label models 1c and 1c').

Research Question 2

We address RQ2 using simple descriptive statistics. The school response variables are described in the Measures section, and Appendices C and D.

Research Question 3

We answer RQ3a and RQ3b using multilevel models. Consider model 3a, in which cohort and subject indices have been suppressed for simplicity. For each construct Z^c (for c = 1,…,5) we fit the following model:

$$Y_{ijT_1} - Y_{ijT_0} = \alpha_j + \Delta_j F_i + \boldsymbol{\beta X_{ij}} + \text{month}_{iT_0} + \beta\text{month\_difference}_i + e_{ij} \ (\text{model } 3a)$$

$$\alpha_j = \delta_a + \phi_0^c Z_j^c + \gamma_j$$

$$\Delta_j = \Delta_\alpha + \phi_1^c Z_j^c + \eta_j$$

Where:

- F_i is a binary indicator of FSM6 status for pupil i

---

[31] In cases where a very small number of pupils fall in a particular month, we aggregate these fixed effects into longer time periods.
[32] We perform analogous analysis for $\widehat{\Delta}_{T_2}^{\text{k}}$ and $\widehat{\Delta}_{T_1g}^{\text{k}}$.

- δ_a is the mean change in Y for non-FSM6 pupils in a typical school;
- $\Delta_a + \delta_a$ is the mean gain score for FSM pupils in a typical school;
- $\boldsymbol{X_{ij}}$ is a matrix of student and school controls (at the pupil level: gender and EAL; at the school level: percentage of pupils achieving expected standard in reading, writing, and maths in 2019, percentage of pupils achieving higher standard in reading, writing, and maths in 2019, inspection ratings (as factor), percentage of pupils who were FSM6 in 2019);
- Z_j^c is a zero-centered school construct describing school responses from March to May 2020 (see Measures section for more detail);
- ϕ_0^c measures the association between school-level construct Z^c and the mean change in Y for non-FSM students;
- ϕ_1^c measures the association between school-level construct Z^c and the mean change in Y for non-FSM students (i.e. the association between Z^c and the FSM gap);
- $e_{ij} \sim N(0, \sigma^2)$;
- γ_j and η_j are school-level random effects, where:

$$\begin{pmatrix} \gamma_j \\ \eta_j \end{pmatrix} \sim N\left(\mathbf{0}, \begin{matrix} \sigma_\alpha^2 & \sigma_{a\eta} \\ \sigma_{a\eta} & \sigma_{\Delta_j}^2 \end{matrix}\right);$$

- γ_j is the school-effect for school j for non-FSM pupils (i.e. the mean change in Y at school j for non-FSM pupils);
- $\gamma_j + \eta_j$ is the school-effect for school j for FSM pupils (i.e. the change in Y at school j for FSM pupils);
- η_j is the estimand for the differential impact school j has on the change in Y of FSM (compared to non-FSM pupils) after controlling for the variables in $\boldsymbol{X}$

Our analysis focuses on point estimates of ϕ_1 along with 95% confidence intervals, using profile likelihood.

The analysis for RQ3b uses the same setup, but with two changes:
- We focus on variables from Teacher Survey 2 ($\tilde{Z}^c$) rather than Teacher Survey 1;
- the LHS variable is the change in achievement from September/October 2020 (T_1) to December 2020 (T_2)

$$Y_{ijT_2} - Y_{ijT_1} = \alpha_j + \Delta_j F_i' + \boldsymbol{\beta X_{ij}} + \text{month}_{iT_0} + \beta\text{month\_difference}_i + e_{ij} \text{ (model 3b)}$$

$$\alpha_j = \delta_a + \phi_0^c \tilde{Z}_j^c + \gamma_j$$

$$\Delta_j = \Delta_\alpha + \phi_1^c \tilde{Z}_j^c + \eta_j$$

Deviations and extensions from the study plan

1. The study plan specifies that we will fit model 1a allowing for the parameter estimates to vary by year group. To simplify estimation and interpretation of this 'by year group' analysis, we fit model 1a separately for each year group (g) and subject (k)—for a total of 10 sets of estimates.[33] The same applies to models 1b and 1c.
2. In addition to estimating changes in disadvantage gaps between Autumn 2019 and Summer 2020 (Δ_{T_1}) and changes from Summer 2020 to Autumn 2020 (Δ_{T_2}) we also estimate a 'summary' change covering the entire study period: Autumn 2019 to Autumn 2020.
3. We present 'summary' results averaging across Years 2–6. As noted above, this involved estimating $\hat{\Delta}_{T_X}^{\text{k}}$—an average change in disadvantage gaps across Years 2 to 6 for outcome k, in time period T_X—in

[33] This deviation applies to RQ1a and RQ1b.

addition to the year-level-specific estimates $\widehat{\Delta}^{k}_{T x g}$. See Appendix E for details of likelihood ratio tests. We present $\widehat{\Delta}^{k}_{T x}$ estimates in Table 10.

4. The study plan states that we would adjust for the timing of tests using the date of test administration (in RQ1a, RQ1b, RQ3a, RQ3b. Unfortunately, one of our sources of date information only reported information on the month in which pupils took tests. To make our data consistent across sources we therefore used 'months' throughout. In the robustness section we provide evidence that the timing of the tests is not an important factor in the evolution of disadvantage gaps (Figure 12 presents the results for reading and maths averaged across Years 2 to 6. Including time controls barely changes our estimates.). As such, we remove these controls for our analysis of RQ3. The associations we estimate are very insensitive to this change. A likely explanation for this lack of sensitivity is that our analysis focuses on a *difference* (i.e. the disadvantage gap) rather than the *level* of attainment, which is more likely to be systematically influenced by the date of test administration.
5. In RQ3, we did not include 'ethnic background' as a covariate. The reason was that the number of predictors introduced by this categorical variable made models unstable.
6. In RQ3, we present the 'main effects' of each school construct (ϕ_0^c) in addition to the main focus of our analysis, namely the association between school responses and disadvantage gaps. We believe this is of substantive interest.
7. In the analysis plan for RQ3, models 3a and 3b specified zero-center the FSM6 indicator. Upon reflection, this made parameter interpretation more difficult, so we left the FSM6 indicator as binary.
8. There were several deviations to the Robustness section of our study plan:
    a. The study plan said that we would examine whether our results were robust to including individual-level controls for absences. Unfortunately, technical problems prevented us from collecting absence data from schools within the timeframe of the project.
    b. The plan also said that we would examine the effect of how school responses Z and $\tilde{Z}$ were scaled. However, we ultimately decided not to conduct these sensitivity checks. The survey data we gathered ended up being less detailed than we anticipated when we wrote the study plan. In particular, we had thought there may be a large number of questions for teachers. But, in an effort to minimise the burden on teachers, we ultimately had quite short surveys (especially for survey two). This being the case, we did not think it was justified to use IRT on constructs that rely on one or two questions.
    c. Finally, we did not run the 'repeated measures models' from the study plan. We wanted to limit the already large number of analyses we were presenting and believe that the models we present are simpler and preferable to more complex models.

Research findings

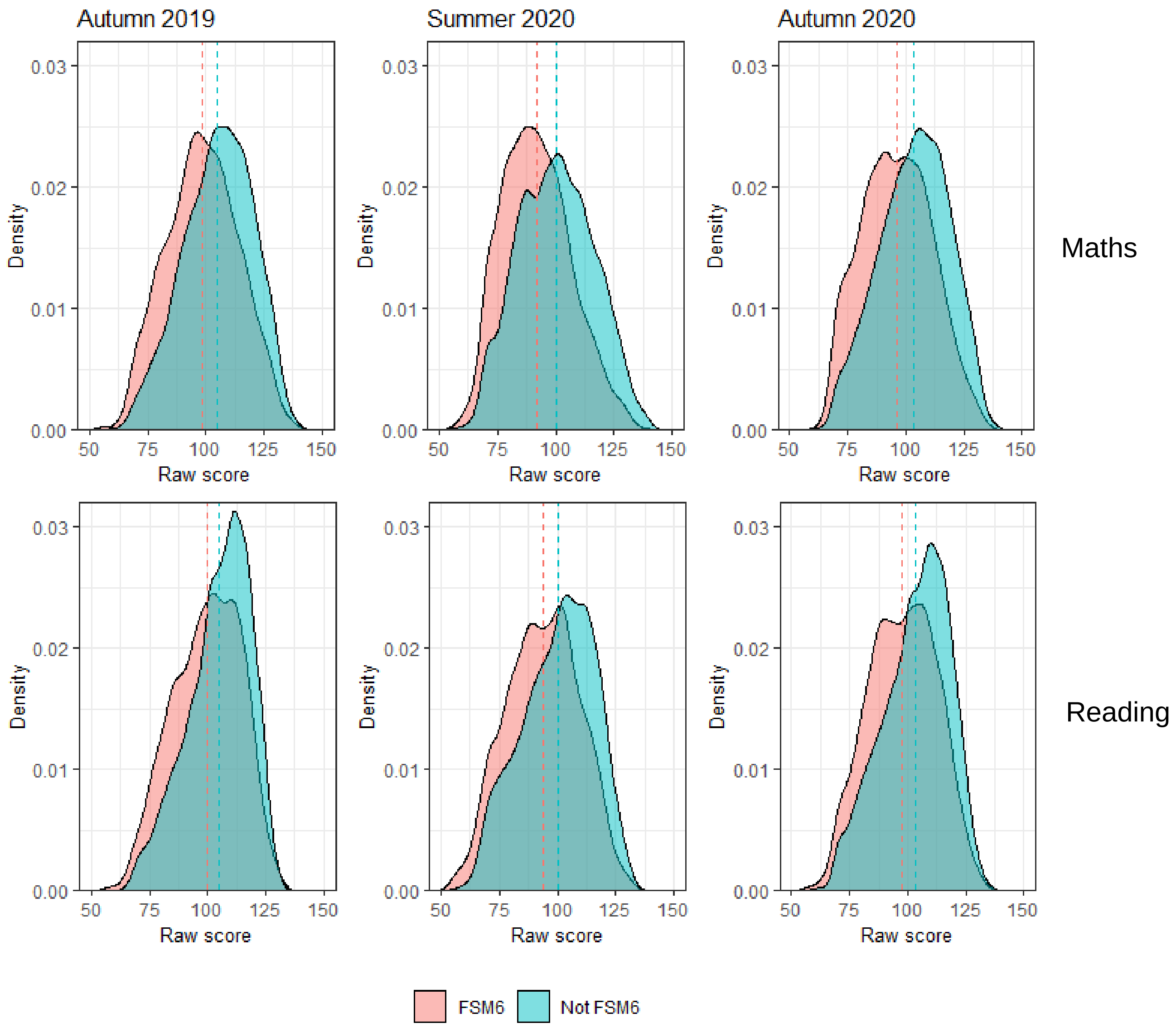
Results

Research Question 1

Distributions of attainment data

We begin by plotting the distributions of RS Assessment test data, for the three assessment periods (Figure 5). These distributions cover Years 2–6 and do not appear to suffer from ceiling or floor effects. In all three periods we observe a gap between FSM6 pupils and their peers (indicated by the difference between the dotted lines on each plot, which represent means).

Figure 5: Distribution of attainment scores (Y^{Norm})



Notes: 'Raw score' is Y^{Norm}, described in the Measures section. The sample we use throughout is the RQ1 responder sample.

Estimating disadvantage gaps at the start of the study: Autumn 2019

To contextualise changes in the magnitude of disadvantage gaps, Table 8 presents the level of disadvantage gaps before covid disruptions, using equation (1). The scale is effect size units. In our sample, the average disadvantage gap is 0.39 ES in reading and 0.44 ES in maths. Estimates from the National Pupil Database suggest that, across primary school, the average disadvantage gap is around 0.50 ES units (Thomson, 2021). This suggests that the gaps in our sample are similar, but narrower, to disadvantage gaps in the broader population.

Table 8: Estimated disadvantage gap in Autumn 2019 ($\hat{G}_{Aut19}$) of RQ1 responder sample

	Reading (s.e.)	**Maths (s.e.)**
Year 2	0.41 (0.06)	0.36 (0.06)
Year 3	0.44 (0.07)	0.41 (0.06)
Year 4	0.35 (0.05)	0.48 (0.05)
Year 5	0.31 (0.05)	0.44 (0.05)
Year 6	0.42 (0.05)	0.51 (0.05)
Overall average (Years 2–6)	**0.39**	**0.44**

Notes: throughout this study year levels are defined by the 2020–2021 academic year. 'Year 3' in this table refers to the cohort of children who were in Year 2 in 2019–2020. Standard errors are in parentheses. The definition of 'disadvantage gap' is given in equation (1). The scale is effect size (ES) units. For sample sizes of the RQ1 responder sample, see Figure 2.

Mean attainment by FSM6, during the three assessment windows

Next, we present information about mean attainment, for FSM and non-FSM students, in maths and reading (Table 9).[34] These simple means represent averages across year levels. They provide an initial sense of how disadvantage gaps have changed over time, without using statistical models We use standardised scores, as described in the Measures section. Note that these scores differ from the raw attainment data we received from RS Assessment, presented in Figure 5.

Table 9: Mean standardised scores in reading ($Y_{ij}^{Reading}$) and maths (Y_{ij}^{Maths}) by FSM status, in RQ1 analysis sample

	Reading			**Maths**		
	Not FSM6 (s.e.)	FSM6 (s.e.)	Gap	Not FSM6 (s.e.)	FSM6 (s.e.)	Gap
Autumn 2019	0.13 (0.01)	−0.26 (0.02)	0.39	0.13 (0.01)	−0.31 (0.02)	0.44
Summer 2020	0.21 (0.01)	−0.22 (0.02)	0.43	0.24 (0.01)	−0.29 (0.02)	0.53
Autumn 2020	0.17 (0.01)	−0.25 (0.02)	0.42	0.19 (0.01)	−0.32 (0.02)	0.51
$\boldsymbol{n}_{\textbf{pupils}}$	6043	2416		6172	2415	

Notes: standard errors for mean estimates are in parentheses. All numbers are in ES units.

Detailed analysis for RQ1a and RQ1b

Next, we present results from statistical models that directly estimate changes in disadvantage gaps. We start with the most aggregated level: changes in the disadvantage gap for reading and maths, averaged across Years 2 to 6. The results are presented in Table 10, which has three columns. The first column presents estimates of gap changes during the period from Autumn 2019 to September 2020. During this 'initial closure period', point estimates for the change in the disadvantage gap are negative (implying a widening); however, there is considerable uncertainty. For maths, the point estimate is a 0.065 ES unit widening, with a 95% confidence interval of [−0.098 to −0.031]. For reading, the point estimate suggests a smaller widening of 0.010 ES units, with a 95% confidence interval of [−0.048, 0.028].

The second column presents estimates from the 're-opening period' during which schools across the country were generally open for face-to-face instruction, albeit with widespread disruption due to covid. Here we see fewer signs

[34] See the Measures section for definitions of the standardised attainment outcomes.

of changes in economic gaps, with point estimates close to zero for both reading and maths. In short, during the first term in which students went back to school, we see no clear evidence of gaps closing, or widening further.

The final column is a summary and encompasses columns 1 and 2. Over the year from November/December 2019 to November/December 2020 we find evidence that the disadvantage gap for primary maths widened. Our point estimate is −0.075 ES units [95% CI −0.108, −0.043] which represents a 17% increase in the pre-covid gap.[35] A plausible range is a widening of between 10% and 24%. Our estimate for the change in the disadvantage gap in reading is −0.003 with a 95% CI of [−0.041, 0.036]. While there is uncertainty, this suggests that the gap in reading is similar to pre-covid levels.

Table 10: Estimated average change in the disadvantage gap

	Autumn 2019–September 2020 (Closure period)	September 2020–Autumn 2020 (Re-opening period)	Autumn 2019–Autumn 2020 (Summary)
Maths **[95% CI]**	−0.065 [−0.098, −0.031]	−0.008 [−0.037, 0.021]	−0.075 [−0.108, −0.043]
Reading **[95% CI]**	−0.010 [−0.048, 0.028]	0.010 [−0.024, 0.043]	−0.003 [−0.041, 0.036]

Notes: full regression results underlying these plots are presented in Appendix F. All analyses rely on the RQ1 full responder sample, described in Figure 2, and compare to broader populations in the section on Comparing analysis samples.

Would we expect gaps to grow within a cohort?

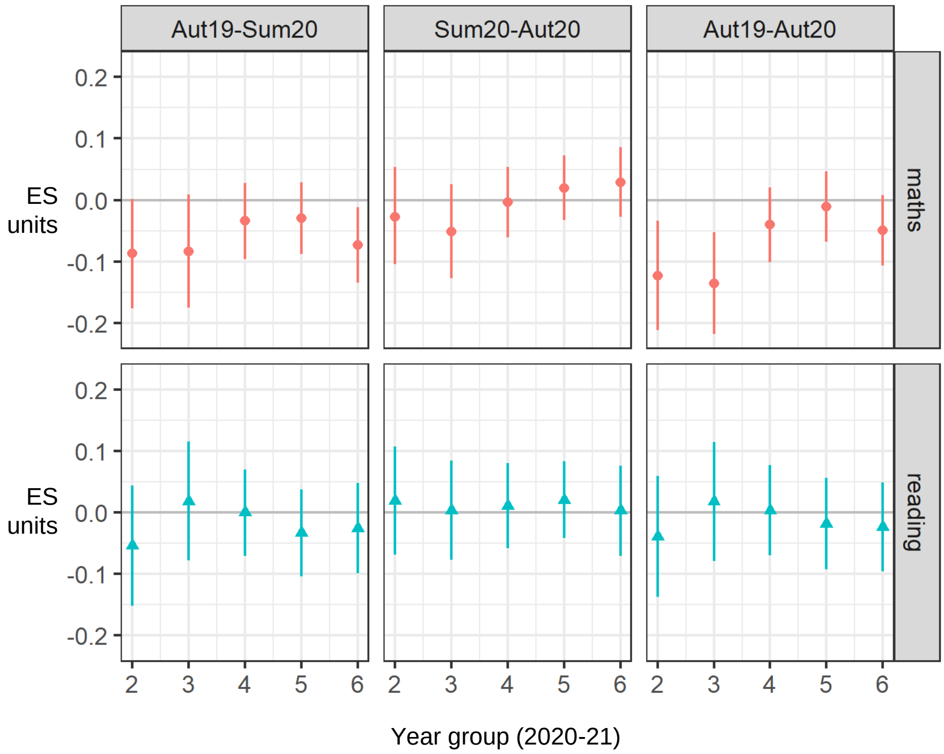
Our analysis tracks how disadvantage gaps have changed over a year, for a constant group of students (spanning Years 2–6). It is worth considering how these direct estimates of gaps typically evolve over time in the absence of covid. Such an analysis was recently published using National Pupil Database (NPD) data (Thomson, 2021). The analysis examines two separate cohorts in the NPD: one who finished Key Stage 2 in 2019, and another who finished Key Stage 2 in 2018. The results differ slightly by cohort, but in both cases the magnitude of the average annual change in the disadvantage gap was smaller than 0.005 ES units. In other words, in the absence of covid-19, gaps within a cohort change extremely gradually from year to year. This suggests that the changes we present in Table 10 are substantial in comparison, and are a result of covid-19 disruptions.

Results by year level

Figure 6 presents results disaggregated by year group. The largest point estimates of gap widening are among maths outcomes for students currently in Years 2 and 3. This echoes other work suggesting that younger children from disadvantaged backgrounds may have been particularly affected by covid disruptions (Curriculum Associates, 2020; Juniper Education, 2020). That said, we do not have enough data to draw strong conclusions about whether gaps have grown more for younger children. As discussed in Appendix E, we fail to reject the hypothesis that there are no effects of year level on changes in attainment gaps for reading and maths. While there do appear to be subject differences (between reading and maths) in our data, we urge caution against making overly specific distinctions about age–subject combinations who have been negatively affected by covid-19.

[35] The average disadvantage gap for maths in our sample at the Autumn 2019 baseline was 0.44 ES units as per Table 8. To calculate the percentage change, we divided 0.075 by 0.44.

Figure 6: Results from RQ1a and RQ1b by year level



Notes: negative estimates indicate a widening of gaps. Full regression results underlying these plots are presented in Appendix F. All analyses rely on the RQ1 full responder sample, described in Figure 2, and compared to broader populations in the section on Comparing analysis samples to broader samples/populations.

Research Question 2

Teacher Survey 1: School Responses during the 'initial closure period' from March–May 2020

We used five variables to describe school responses during the first period of national lockdown (see Table 7 for overview and Appendix C for details). Figure 7 provides a breakdown of how responses varied across teachers and schools. At the teacher level there was a very even spread of responses in all categories, as is clear from tables of teacher counts. There were two minor exceptions: relatively few teachers reported having a daily instructed timetable (80 out of 539) or live lessons (66 out of 539). Overall, the teacher counts show that there was considerable variability in the response to the initial lockdown.

There was also variability at the school level, demonstrated by the histograms in Figure 7 which plot school-level means. That said, two of the variables (timetabling and phoning students) had a clear, modal response—indicated by the respective spikes in the histograms for those variables—which resulted in less variability at the school level. This limits our statistical power in RQ3.

Teacher Survey 2

In this section we focus on the period from September 2020 to December 2020, during which most students returned to schools. As described in the Measures section, we summarise the Teacher Survey from this period using five variables (see Appendix D for details). Figure 8 is analogous to Figure 7 and provides a breakdown of how responses varied in the second Teacher Survey. Again, we note that responses at the teacher level are fairly evenly spread. The 'extra time' variable is an exception: most teachers reported that schools had not pursued this strategy in the Autumn term (341 of 454) and those that did report some extra time almost all reported learning during lunchtime (rather than longer days, or being open in the holidays). At the school level, the five variables all have some variation. In particular, schools showed a varied response to the questions about responding to lost learning by reducing the curriculum and/or using small group interventions. There was less variation in providing 'extra time' (this was a relatively rare response). Similarly, only a minority of schools provided video/streamed classes to pupils who were absent (as opposed to whole classes who were isolating).

Figure 7: Descripton of responses to Teacher Survey 1—tables of teacher counts and histograms of school means

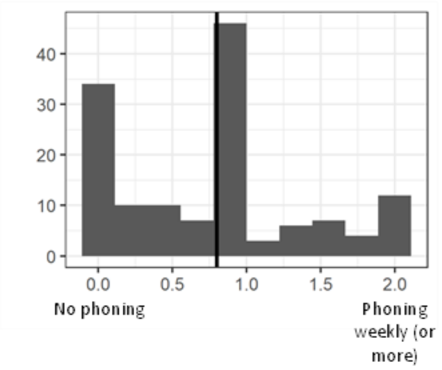
Histogram of school means

Phoning students

Did teachers speak to students (not parents) directly on the phone? 0 = not at all during the half-term; 1= either one or twice during the half term; 2 = about once a week or more.

Teacher responses (counts)

0 (no phoning)	1	2 (weekly or more)
200	216	123

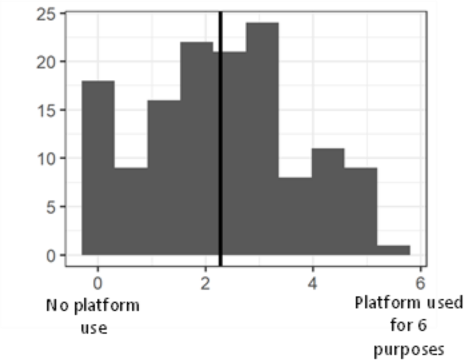


Number of purposes for online Platforms *(Platform)*

Did schools use a platform: to set work (+1); to receive work (+1); to share examples of student work (+1); for daily registration (+1); for text chat interactions (+1); for live audio/video chat interactions(+1)? One point for each, for a maximum of 6 total points.

Teacher responses (counts)

0 uses	1	2	3	4	5	6 uses
91	80	94	130	92	41	11

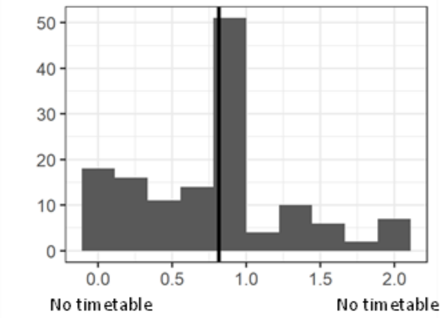


Timetabling approach *(Timetable)*

What timetabling approach did schools have?
0 = no timetable; 1 = a suggested timetable; 2 = instructed daily timetable.

Teacher responses (counts)

0 (no timetable)	1	2 (instructed daily timetable)
179	280	80

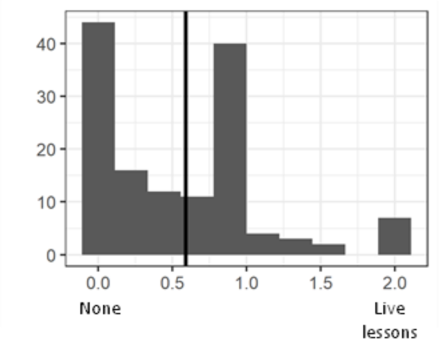


Videos/Live Lessons *(Vid/live)*

Did teachers have pre-recorded video lessons (scored 1) or live lessons (scored 2)? If neither, the variable is coded as zero.

Teacher responses (counts)

0 (none)	1 (video)	2 (live)
282	191	66



Frequency of work submission *(Work submission)*

How often did teachers expect work to be submitted? 0 = no recommendation about work submission; 1 = suggested work submission once a week or fortnight; 2 = several times a week; 3 = every day.

Teacher responses (counts)

0 (no requirement)	1	2	3 (daily)
185	104	84	166

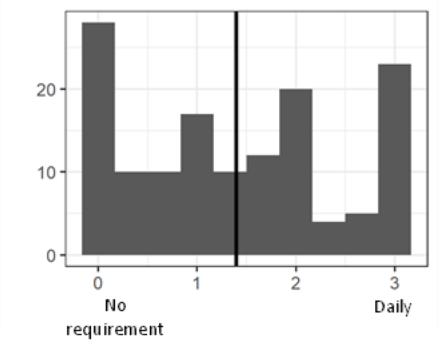


Figure 8: Description of responses to teacher survey 2: tables of teacher counts and histograms of school means

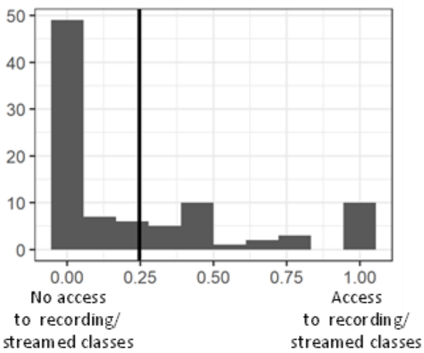
Absence provision for *individuals*

Did schools provide access to video/streaming of classes when individual children were absent? 0 = no access to any video/streaming of class; 1=video/streaming class

Teacher responses (counts)

0 (no access to video/stream)	1 (video/streamed lessons)
346	108

Histogram of school means

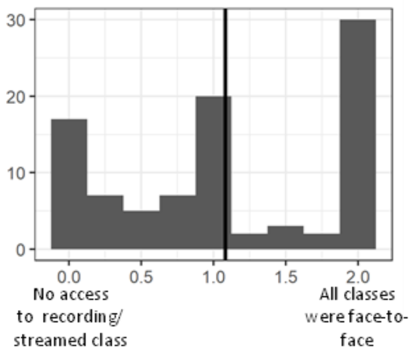


Absence provision for *classes*

Did schools provide access to video/streaming of classes when who classes needed to isolate? 0=no access to any video/streaming; 1=video/streaming; 2=no class isolation needed

Teacher responses (counts)

0 (no video lessons)	1 (recording/stream)	2 (face-to-face)
169	131	154

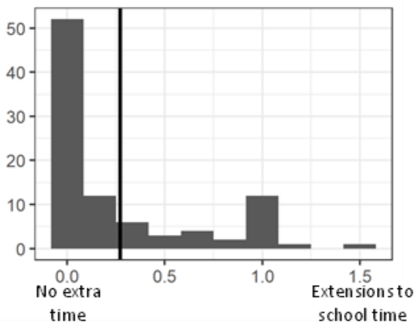


Extra time

Did schools extend the school day (+1); run extra lunchtime learning activities in order to compensate for lost learning (+1); remain open during holidays (+1)?

Teacher responses (counts)

0 (no extra time)	1	2	3 (multiple extensions to school time)
341	104	6	3

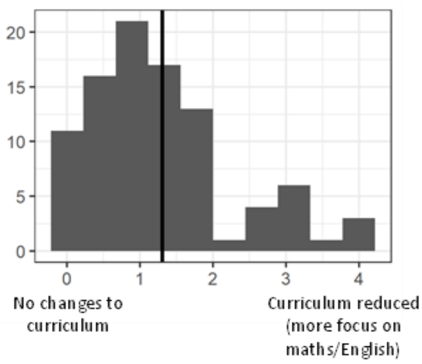


Reduced curriculum

Did schools reduce the curriculum, and focus more time on English and maths in the autumn 2020 term? Composite measure of "removing items from the curriculum" (+1) and "how many extra minutes per day were spent on English and maths (15 mins=1, 30 mins=2, 45+ mins=3, zero mins=0)

Teacher responses (counts)

0 (no change)	1	2	3	4 (reduced curriculum)
171	102	93	51	37

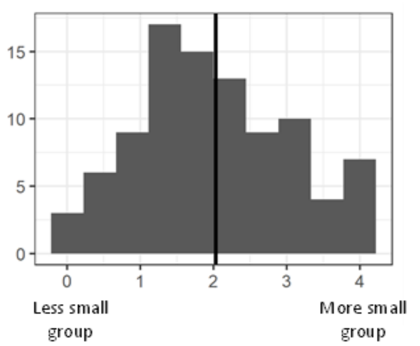


Small group interventions

Did schools make more/less use of small group and individual interventions (e.g. run by TAs)? 0=fewer interventions; 1=same number as normal; 2 to 4 is more than normal, either during lessons, or outside lessons, or both

Teacher responses (counts)

0 (less small group)	1	2	3	4 (more small group)
73	102	90	117	72



Research Question 3

Research question 3 synthesises the analysis from RQ1 and RQ2. We explore whether there are any associations between the school responses captured in our teacher surveys and changes in attainment gaps.

Analysis of Teacher Survey 1 and outcomes from before and after the 'initial closure period'

We begin by estimating the association between relative changes in attainment over the initial closure period (Autumn 2019 to Summer 2020) and the five school response variables from Teacher Survey 1 ('phoning students'; 'use of technology platforms'; 'timetabling'; 'video/live lessons'; 'frequency of work submission').36 The left panel of Figure 9 presents these associations, conditional on a set of pupil and other school covariates.[37] We fail to find evidence of conditional association between relative change in attainment and any of the five response variables. None of the point estimates are greater than 0.06 ES in magnitude, and all 95% confidence intervals contain zero. Similarly, we find no evidence that the raw associations have large, non-zero associations.[38]

Given our focus on inequality we now turn to the conditional association between the school responses from Teacher Survey 1, and **changes in disadvantage gaps** over the initial closure period (Autumn 2019 to Summer 2020). These conditional associations are presented in the right panel of Figure 9.[39] Once again we fail to find clear evidence of non-zero associations. All point estimates are smaller than 0.03 ES units, with the exception of the association between timetabling and reading, with an estimate of 0.063 ES units and a CI of [0.001,0.125].[40] However, in the absence of a clear a priori hypothesis about why this particular practice should be associated with reading and not mathematics, and the exploratory nature of the study with multiple comparisons, we are cautious about interpreting one statistically significant result as anything we would expect to be reproducible.

It is also worth noting that none of the school responses were randomised, so the estimated associations we present rely entirely on the conditional independence assumption for their causal credibility. While there is growing evidence from within-study comparisons that this assumption is often plausible in education settings—and that typical levels of selection bias in education may be smaller than previously thought (for example, Cook *et al.*, 2008; Weidmann and Miratrix, 2020; Wong *et al.*, 2017)—given that we are not conditioning on a student-level measure of attainment, we place limited credence in any causal interpretation of our results.

---

[36] Plots illustrating the distribution of change in attainment are presented in Appendix I.

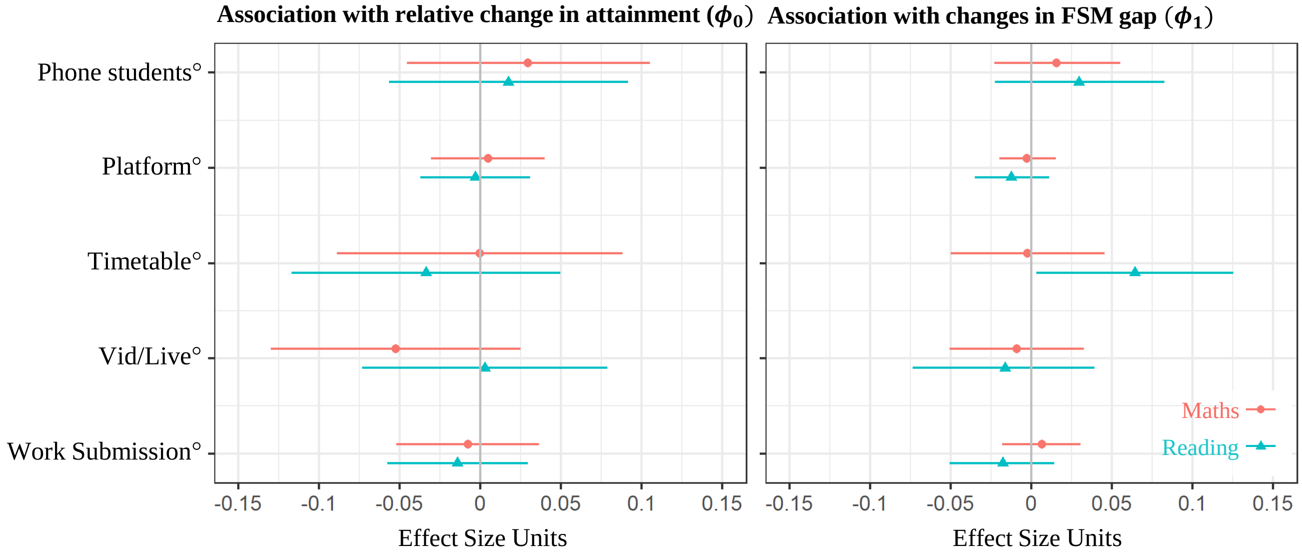
[37] These are the coefficients from model 3a. At the pupil level we condition on gender and EAL status. At the school level we condition on the percentage of pupils achieving the expected standard in reading, writing, and maths in 2019; the percentage of pupils achieving higher standard in reading, writing, and maths in 2019; Ofsted results; the percentage of pupils who were FSM6 in 2019.

[38] None of the raw correlations between school response and the change in relative achievement ($Y_{ijT_1} - Y_{ijT_0}$) are greater than 0.06 in magnitude. There is no evidence to suggest any of these associations are different from zero. The lack of association in the left panel of Figure 9 does not appear to be a function of conditioning or modelling.

[39] These are ϕ_1 coefficients from model 3a.

[40] We thank an anonymous reviewer who pointed out the following *post hoc* hypothesis to explain the reading–maths difference in timetabling. Large numbers of schools use Accelerated Reader and use a timetabled approach (for example, 20 mins per day). Such interventions are more prevalent for literacy than maths. While debate exists over the wider efficacy of such programmes, this is a potential explanation for effect on reading and a higher level of timetabling, namely that the structure was associated with the involvement of a programme which revealed itself to be helpful during lockdown.

Figure 9: Associations between school responses during the initial closure period and 'change in attainment' (left panel) and 'change in disadvantage gap' (right panel)



Notes: both panels are based on model 3a. The dependent variable is $Y_{ijT_1} - Y_{ijT_0}$, the estimated change in learning from Autumn 2019 to Summer 2020. °School response variables are defined in the Measures section (see Table 7 for an overview) and have been centred, to have a mean of zero. We have retained the original scaling to make the results as interpretable as possible. The sample here is the analysis sample for RQ3a, described in Figure 3 and consisting of 90 schools for reading and 88 schools for maths. Conditional associations between school response variables and changes in learning are presented in the left panel and represent $\hat{\phi}_0$ parameters from model 3a. Condition associations between school response variables and changes in disadvantage gaps are presented in the right panel and represent $\hat{\phi}_1$ from model 3a. In both cases we condition on gender and EAL status at the pupil level and the following school-level variables: the percentage of pupils achieving expected standard in reading, writing, and maths in 2019; the percentage of pupils achieving higher standard in reading, writing, and maths in 2019; Ofsted results; the percentage of pupils who were FSM6 in 2019. 95% confidence intervals come from profile likelihood. Full regression results are in Appendix G.

Analysis of Teacher Survey 2 and changes in outcomes across the Autumn 2020 term

Our analysis here follows the same pattern as our analysis of Teacher Survey 1. We start by examining the associations between relative **changes in attainment** during the 're-opening period' (Autumn term of 2020) and the five school response variables from Teacher Survey 2 ('absence provision for *individuals*'; 'absence provision for *classes*'; 'extra time'; 'reduced curriculum'; and 'small group interventions'). The left panel of Figure 10 presents these associations, conditional on a set of pupil and other school covariates.[41] We fail to find evidence of conditional association between relative change in attainment and any of the five response variables. All 95% confidence intervals contain zero. Similarly, we find no evidence that the raw associations are different from zero.[42]

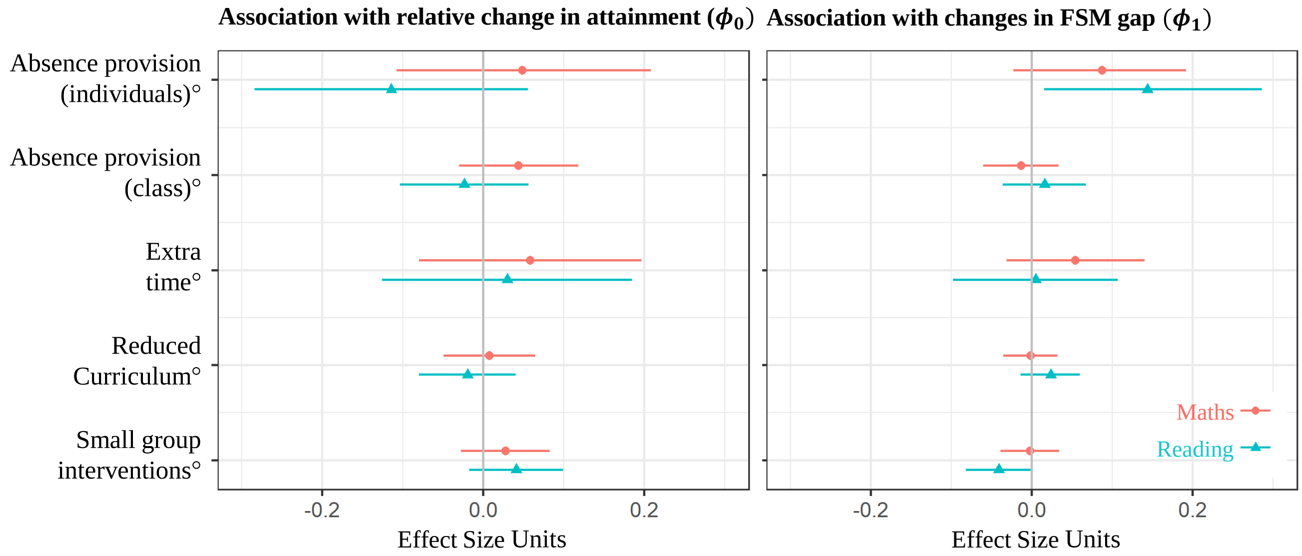
The right-hand panel of Figure 10 presents the association between **changes in disadvantage gaps** over the re-opening period (Autumn 2020 term) and the five school response variables we captured in Teacher Survey 2. Once again, we fail to find clear evidence of non-zero associations. That said, there appears to be some evidence that providing videos or live lessons to absent pupils was associated with a narrowing of disadvantage gaps. In schools where absent pupils had access to video or live recordings of lessons, we estimate that gaps in the Autumn term narrowed in maths by 0.087 [−0.023,0.191] and in reading by 0.143 [0.015, 0.286]. These positive estimates are robust to different sets of covariates in model 3a. However, we note that these are uncertain findings, and could have arisen due to chance. We think it is conceivable that a replication of this

[41] These are the coefficients from model 3b. At the pupil level we condition on gender and EAL status. At the school level we condition on the percentage of pupils achieving expected standard in reading, writing, and maths in 2019; the percentage of pupils achieving higher standard in reading, writing, and maths in 2019; Ofsted results; the percentage of pupils who were FSM6 in 2019.

[42] None of the raw correlations between school response and the change in relative achievement ($Y_{ijT_2} - Y_{ijT_1}$) are greater than 0.08 in magnitude. There is no evidence to suggest any of these associations are different from zero. In short, the lack of association in the left panel of
Figure 10 does not appear to be a function of conditioning or modelling.

study may identify a different school practice as having the strongest association with narrowing disadvantage gaps.

Figure 10: Associations between school responses during the re-opening period and 'change in attainment' (left panel) and 'change in disadvantage gap' (right panel)



Notes: Both panels are based on model 3b. The dependent variable is $Y_{ijT_2} - Y_{ijT_1}$, the estimated change in learning from Summer 2020 to Autumn 2020. °School response variables are defined in the Measures section (see **Error! Reference source not found.** for an overview) and have been centred, to have a mean of zero. We have retained the original scaling to make the results as interpretable as possible. The sample here is the 'analysis sample for RQ3b' initially discussed in Figure 4 and consisting of 58 schools. Conditional associations between school response variables and changes in learning are presented in the left panel and represent $\hat{\phi}_0$ parameters from model 3b. Conditional associations between school response variables and changes in disadvantage gaps are presented in the right panel and represent $\hat{\phi}_1$ from model 3b. In both cases we condition on gender and EAL status at the pupil level and the following school-level variables: the percentage of pupils achieving expected standard in reading, writing, and maths in 2019; the percentage of pupils achieving higher standard in reading, writing, and maths in 2019; Ofsted results; the percentage of pupils who were FSM6 in 2019. Full regression results are in Appendix G.

Looking across all the findings for RQ3, the dominance of null results is arguably surprising. We examined what we believed were strong candidates for school-level variables that could predict changes in attainment and disadvantage gaps. This included resource-intensive efforts, such as providing live-steam videos during the initial lockdown.[43] However, with some tentative exceptions, it is not clear that any of these variables had strong associations with changes in attainment, or its social gradient.

Among possible explanations for these mostly null findings, we note three in particular. First, our measures of 'school' practices are noisy. As noted in the Measures section, teachers within a school sometimes provided different responses. Ideally, our analysis would have been conducted at the class level so that we could directly link teacher reports to pupil attainment. Unfortunately, we were unable to link teacher surveys to specific students, which meant that we had to perform our analysis at the school level.[44] This introduced significant measurement error into our analysis.

Second, for some of the variables there was a lack of variation in school-level responses. In some cases, relatively few schools pursued a particular approach. For example, very few schools reported extending the school day during the Autumn 2020 term (4 out of 93). Similarly, 80% of schools in our sample reported having no live lessons during the initial national lockdown (109 out of 137). This limits our ability to test whether these practices were associated with changes in learning inequalities.

[43] We acknowledge that education evaluations often provide null results, so perhaps we should not be surprised. However, most evaluations study relatively limited interventions affecting a small minority of instruction time. In this case, we examined larger differences (for instance 'receiving online lessons' vs 'no lessons at all'). In other words, we were examining what we believed to be a sharper contrast than is the case in a typical evaluation.
[44] Our data provided no way of knowing which teachers taught which children, and collecting this would have been an additional burden on schools. Moreover, our teacher survey was anonymous.

Third, our examination of practices during the initial lockdown involves tests taken in November/December 2019 and September 2020. The school responses we examined were only operating during a minority of this period. Moreover, this was a tumultuous time: perhaps we should not be surprised that the influence of schools would be overwhelmed by other factors. This reasoning applies somewhat less strongly to our analysis of the Autumn 2020 term, as assessments were taken immediately before and after the period of schooling our survey describes. During the Autumn 2020 term, however, schools were substantially disrupted by covid, and their ability to implement any of these responses would have been limited by the extremely difficult circumstances. This last point, about implementation, is worth emphasising. We were unable to capture any qualitative information about the various school practices we analyse, but we imagine there was substantial variation across schools in how they were implemented. This variation may also have contributed to our null results.

Robustness and further analysis

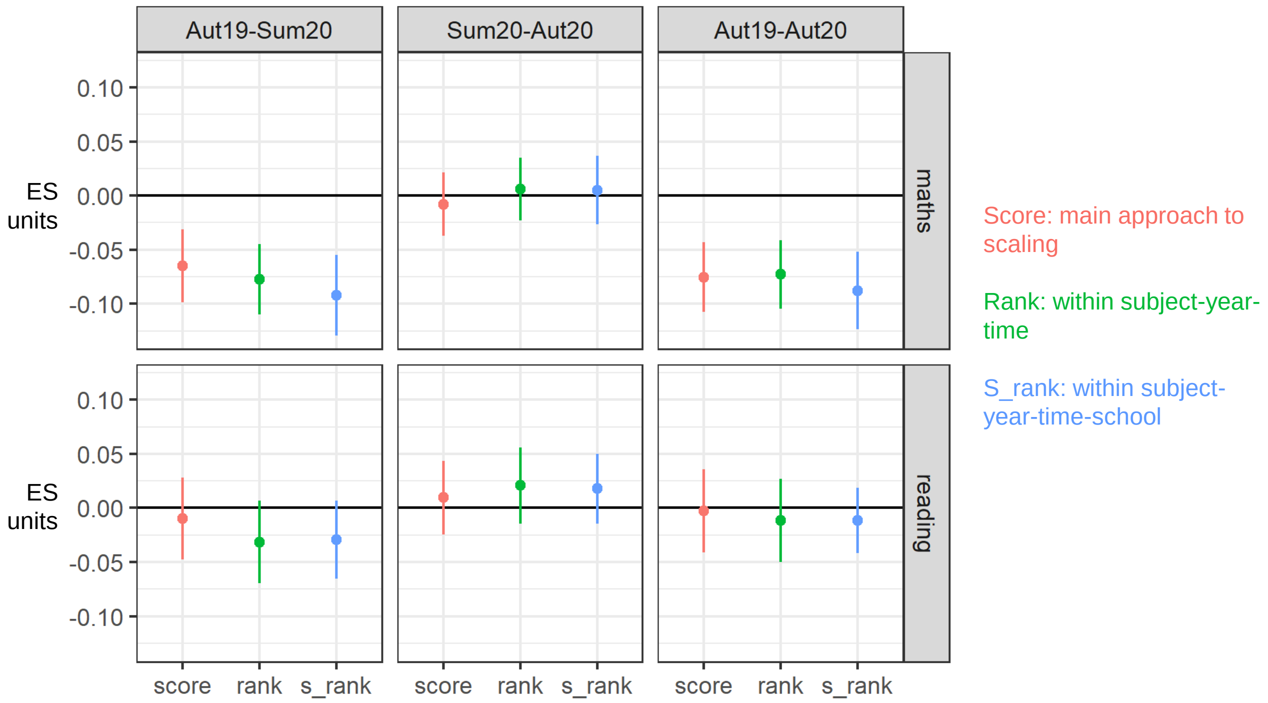
Standardisation/scaling of outcomes

In this section we check whether the results for RQ1 are sensitive to the method of scaling attainment outcomes by comparing our core findings using three different scaling approaches:

1. Standardising scores (this method is represented in the main results);
2. Ranks, defined by assessment/time period/year group;
3. Ranks, defined by assessment/time period/year group/school.[45]

Figure 11 presents the results for these three different methods, illustrating the changes in disadvantage gaps for reading and maths in the three periods (mirroring the structure of Table 10). In all cases we present average impacts across Years 2 to 6. While there are some minor differences—for example, the point estimates for gap widening are slightly larger if we use within-school ranks—the main conclusions are not sensitive to our choice of scaling.

Figure 11: Estimates of changes in disadvantage gaps, using 3 different approaches to scaling



Notes: all estimates average across Years 2 to 6, as described in the Methods section. Point estimates have 95% confidence intervals.

Methods to control for the date of test administration

To assess whether our results are robust to different methods of controlling for test administration date, we compare the results of model 1a with a model that has no time controls (model 1a"):[46]

$$Y_{ijgT_1}^{k} - Y_{ijgT_0}^{k} = \alpha_j + \delta_{T_1g}^{k} + \Delta_{T_1g}^{k} F_i + e_{ijg}^{k} \quad \text{(model } 1a''\text{)}$$

$$\alpha_j \sim N(0, \sigma_\alpha^2)$$

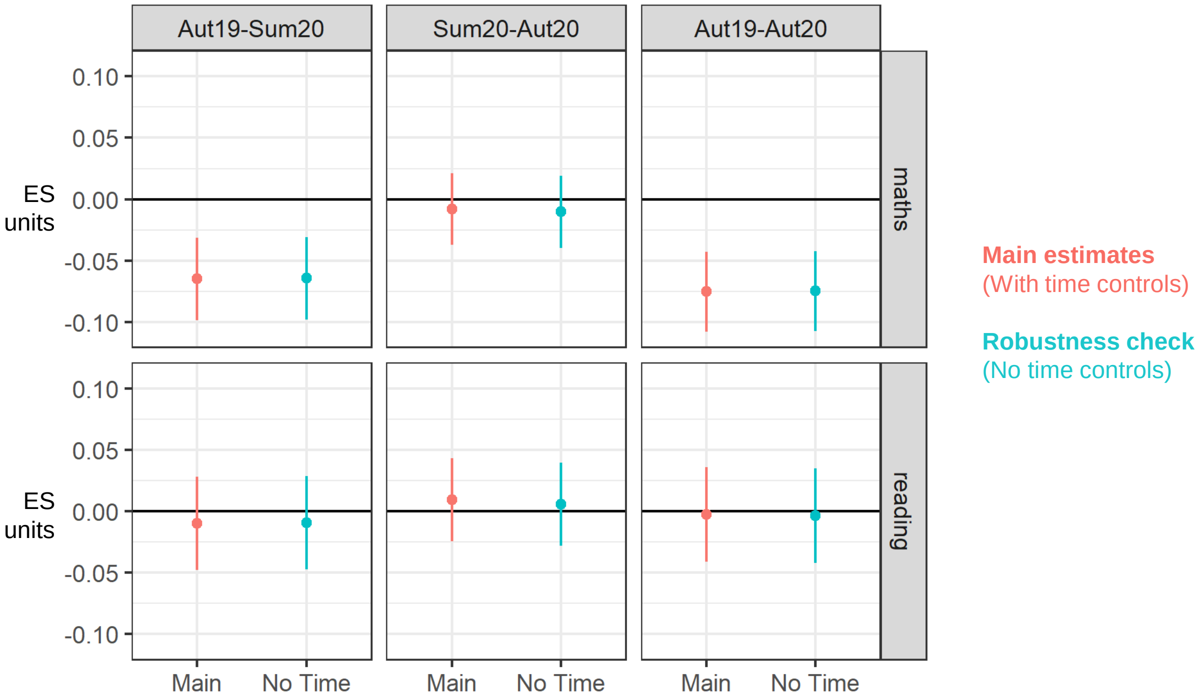
$$e_{ijg}^{k} \sim N(0, \sigma^2)$$

[45] This method was not mentioned in the study plan, but is included to provide a more detailed examination of robustness to scaling.

[46] We fit analogous models for model 1b" and model 1c".

Figure 12 presents the results for reading and maths averaged across Years 2 to 6. Including time controls barely changes our estimates.

Figure 12: Exploring the effect of time controls on $\hat{\Delta}_{T1}$



Notes: main estimates come from the models described in the section on Statistical analysis. Estimates with no time controls come from models 1a'', 1b'' and 1c''. In both cases, we analyse the R1 full responder sample.

Missing data

RQ1 Missingness analysis

As noted in the Participants section, our starting sample for RQ1 was defined by pupils who:

- had a valid assessment in reading or maths in Autumn 2019; and
- were on roll during the Autumn term 2020 and had a disadvantage flag (yes/no) present.

We asked schools to test pupils in each of reading or maths using the Summer 2020 and Autumn 2020 test suites. Based on the response patterns we classified pupils into three groups:[47]

- Full responders: those for whom a Summer 2020 and Autumn 2020 result is observed;
- Partial responders: those for whom a Summer 2020 or Autumn 2020 result is observed (but not both);

Non-responders: those for whom no further test results are observed.Figure 2 illustrates this process. Table 11 provides additional information, broken down into year groups.

Table 11: Response by year group

	Response	Year 2	Year 3	Year 4	Year 5	Year 6	Total
Reading	Full	1317	1237	1928	2217	1760	8459
	Partial	1452	1543	2092	1709	1645	8441
	Non-response	444	434	632	664	1399	3573
	Total	3213	3214	4652	4590	4804	20473

[47] This was done separately for reading and maths.

	Full	1307	1271	1957	2290	1762	8587
	Partial	1546	1545	1763	1425	1655	7934
Maths	Non-response	417	567	702	663	1183	3532
	Total	3270	3383	4422	4378	4600	20053

The principal reason for pupil non-response is school non-response. These are cases where at least two-thirds of pupils in the same national curriculum year were classified as non-responders (effectively attrition at school/year group level). Of the 3,526 non-responders, 2,410 (68%) fall into this category.

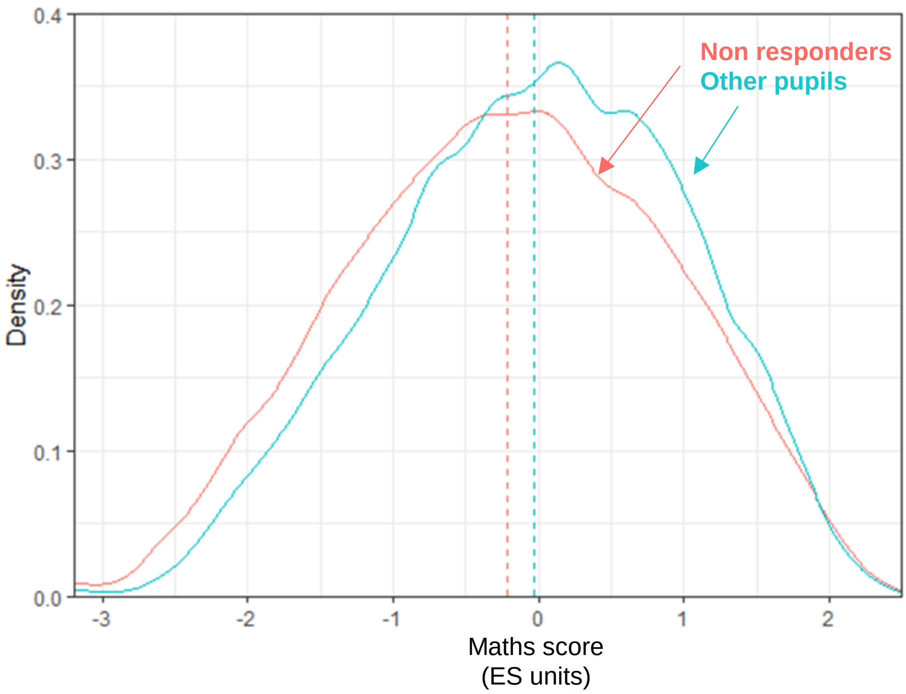
Table 12 below shows that the Autumn 2019 attainment of this group in maths was slightly below average (−0.1 ES units) but was higher than non-responders in schools with <66% non-responders. The table presents the mean Autumn 2019 maths score alongside a version which absorbs the school mean. This suggests that non-responders in schools with a non-response rate of <40% tend not only to be low-attaining but also low-attaining given the school they attend. By contrast, non-responders in schools with a 40% to 66% non-response rate tend to be lower attaining but not especially different to other pupils in the school. This could arise from one class in a year group participating in the project and another one not.

Table 12: Non-responders in maths by school non-response rate

School percentage of non-responders	Autumn 2019 score		
	Actual	Absorbing school mean	Number
<10%	−0.49	−0.34	332
<40%	−0.52	−0.30	438
<66%	−0.29	−0.06	346
≥66%	−0.10	**	2410
All non-responders	−0.21	−0.07	3526

The distribution of Autumn 2019 maths scores for non-responders is compared to other pupils in Figure 13. The non-responder distribution is shifted to the left, suggesting lower baseline performance.

Figure 13: Distribution of prior attainment for non-responders and others for RQ1



Note: dotted mean lines represent means.

Several other forms of non-response can be identified. In *Table 13* we show the number of missing test results in each subject in each term (non-responders plus the partial responders who did not respond in the given term). By far the largest category is school non-response. These are cases where at least two-thirds of pupils in the same national curriculum year at a given school were classified as non-responders.

Table 13: Other sources of non-response for RQ1

	Maths		Reading	
	Summer 2020	Autumn 2020	Summer 2020	Autumn 2020
Leavers	245	374	251	384
Invalid results	78	11	47	12
School non-response	4136	6141	5279	5798
School response pupil non-response	975	658	940	536
Total missing	5434	7184	6517	6730

Reasons for pupil non-response in responding schools include:

- Working below the level of the test;
- Absence;
- Technical issues linking data.

However, we cannot distinguish between the three.

Ultimately, because we are focused on estimates of changes in gaps and we use a consistent sample across time, non-response is not an issue for the internal validity of our estimates—but it does bear on generalisability. With that in mind we turn to the question of whether our 'full responder' sample differs from the broader sample.

Table 14 examines whether observable characteristics differ between our analysis sample ('full responder for RQ1') and the remainder of the sample (partial and non-responders).

Table 14: Observed characteristics of responder and non-responder samples

		Full responders for RQ1	Partial responders for RQ1	Non-responders for RQ1	All pupils
Reading	Autumn 2019 score	0.01	−0.02	−0.12	−0.02
	% disadvantaged	29%	27%	31%	29%
	% EAL	26%	24%	29%	25%
	% SEN	15%	14%	19%	15%
	% female	50%	49%	48%	48%
	% London schools	7%	22%	35%	18%
	% northern schools	22%	25%	20%	23%
	Number of pupils	8470	8433	3570	20,473
Maths	Autumn 2019 score	0.01	−0.06	−0.21	−0.06
	% disadvantaged	28%	29%	36%	30%
	% EAL	23%	22%	29%	24%
	% SEN	16%	15%	20%	16%
	% female	49%	49%	48%	49%
	% London schools	5%	19%	24%	14%
	% northern schools	21%	28%	23%	24%
	Number of pupils	8598	7929	3526	20,053

Non-responders differ from the other categories: they were lower attaining in Autumn 2019 and slightly more likely to be disadvantaged. Pupils attending London schools are disproportionately more likely to be partial and non-responders, probably due to covid-19.

We use logistic regression to see how the characteristics in *Table 14* are associated with 'being excluded from the analysis sample'. Let $P_{ij} = 0$ if student i (from school j) is in the 'partial' or 'non-responder' samples for RQ1 (and $P_{ij} = 1$ if they are in the 'full responder' sample for RQ1):

$$\Pr(P_{ij} = 1) = \text{logit}^{-1}(\boldsymbol{\beta X_{ij}})$$

where $\boldsymbol{X_{ij}}$ are the variables listed in *Table 14*. The results of these regressions are presented in Appendix H. As suggested by Table 3, children who were higher attaining were more likely to remain in our sample and those in London were less likely to be represented. Conditional on other predictors, FSM6 status did not predict missingness for reading, but was negatively associated with missingness for maths.

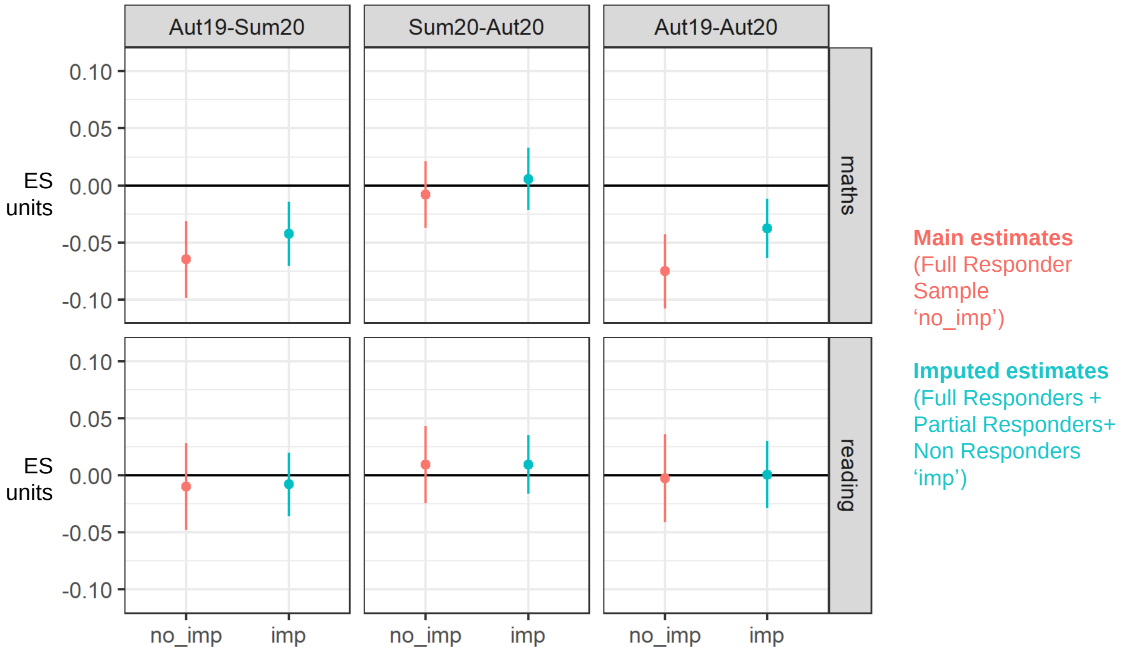
Sensitivity of main RQ1 results to missingness (assuming missing-at-random)

In this section we see if the results presented for RQ1 are robust to multiple imputation using chained equations (MICE). Specifically, we compare estimates from Table 10 to equivalent results when we use MICE to impute outcomes for the Partial and Non-Responder sample. For each combination of year-level/subject/period, we impute 20 datasets.[48] For each dataset, we fit the relevant model (1a', 1b' or 1c') and calculate the relevant $\widehat{\Delta}$ parameter. We then take the average estimate across these 20 imputed datasets and compare the results to those in Table 10. The results of this procedure are presented in Figure 14.The "no_imp" estimates are the original estimates from Table 10, and the "imp" figures are the results from our imputation procedure. The imputation results for reading are almost exactly the same as the main results. For maths, the results with imputed data suggest that gaps may have widened less than our analysis of the full sample suggests. We still find the same broad pattern: i) a widening of maths gaps in the closure period; ii) no noticeable shift in the Autumn term; iii) an overall widening from November/December 2019 to November/December 2020. However, the magnitudes of changes using the imputed estimates are smaller. For the imputed results, the aut19-sum20 estimate is -0.042 (compared to main estimates of -0.065) and for aut19-aut20 the imputed estimate is -0.038, compared to -0.075. We believe that differences between our imputed results and those from the full responder sample may be partly due to 'regression to the mean'. The change scores (e.g. the difference in scores between Autumn 2019 and Summer 2020) are negatively correlated with prior attainment (Autumn 2019 score in this example) and hence exhibit regression to the mean.[49] Disadvantaged pupils with missing post-covid assessment data are disproportionately located in the lower end of the prior attainment range. When we impute we tend to impute larger change scores for disadvantaged pupils than for other pupils. As the percentage of disadvantaged pupils is slightly higher in the non-responder sample for maths, this may partly account for difference between imputed and non-imputed results. However, we emphasize that despite this effect, all confidence intervals of imputed and non-imputed results are overlapping.

[48] For example, for model 1a we impute 20 datasets for the cohort of children entering year 2 in 2020-21, who have missing maths data in the Summer 2020 period. Similarly, we impute 20 maths datasets for the cohort entering year 3, year 4 and so on. Note that this number of dataset imputations is a deviation from our analysis plan, which suggested we would follow the rule-of-thumb recommended by White et al. (2011), and have the same number of imputed datasets as the average percentage of missingness (at student level). This proved to be very time consuming, so instead we present results that are averaged across 20 imputed data samples. The robustness check is not sensitive to this choice.

[49] Correlations are -0.42 for reading and -0.34 for maths.

Figure 14: Comparing main results with imputed results



RQ3 Missingness analysis

This section examines the robustness of RQ3 to missingness and mirrors the analysis examining RQ1 missingness. We start by looking at whether or not there are differences in the observable characteristics between our full teacher survey sample (RQ2) and those for whom we have attainment data. These differences are presented and discussed in Table 5 and

Table 6. To recap: it does not appear as though the schools we analyse in RQ3 (for whom we have overlapping data from teacher surveys and attainment) were unusual in terms of their response to covid disruptions.

We examine this more formally using logistic regression. Let $P'_{ij} = 0$ if student i is in school j that has data for RQ2 (teacher sample) but lacks outcome information for analysis in RQ3 (and $P'_{ij} = 1$ for students who are included in RQ3 analysis). We fit the following model:

$$\Pr\left(P'_{ij} = 1\right) = \text{logit}^{-1}(\boldsymbol{\beta X_{ij}})$$

where $\boldsymbol{X_{ij}}$ are the variables listed in *Table 14*.[50] The results of these regressions are presented in Appendix H. They suggest that, conditional on other observed characteristics, FSM6 students were more likely to attrit (i.e. not have post-covid attainment data, despite attending schools represented in the Teacher Surveys) than their peers. There were no other clear patterns in the data.

Sensitivity of main RQ3 results to missingness (assuming missing-at-random)

Last, we examine the impact of imputing missing values, conditional on observed characteristics, using MICE. Specifically, we compare estimates presented in Figure 9 and Figure 10 to equivalent results when we use MICE to impute outcomes for students who were represented in the teacher survey, but lacked post-covid outcome data.[51] For each analysis (RQ3a and RQ3b) we impute 20 datasets.[52] For each dataset, we fit the

[50] We removed London and northern indicators, as these were very rare in our missingness model and led to unstable estimates.
[51] For RQ3a we were concerned with missing outcome data for September 2020 ('Summer 2020'); for RQ3b, the relevant missingness was for November/December 2020 ('Autumn 2020').
[52] For example, for model 1a we impute 20 datasets for the cohort of children entering Year 2 in 2020–2021, who have missing maths data in the Summer 2020 period. Similarly, we impute 20 maths datasets for the cohort entering Year 3, Year 4, and so on. Note that this number of dataset imputations is a deviation from our analysis plan, which suggested we would follow

relevant models (3a or 3b) and calculate $\hat{\phi}_0$ and $\hat{\phi}_1$ for the school response variables ($\boldsymbol{Z}$ or $\widetilde{\boldsymbol{Z}}$). We then take the average estimate across these 20 imputed datasets and compare the results to those in the results section (Figure 9 and Figure 10). For RQ3a, the results of this procedure are presented in Figure 15. The top row (circles) shows the new estimates, using imputed data. The bottom row (triangles) shows the main estimates. The imputation procedure produces results that are extremely similar. The same is true for the RQ3b, as shown in Figure 16.

Figure 15: Comparison of imputed and main results for RQ3a

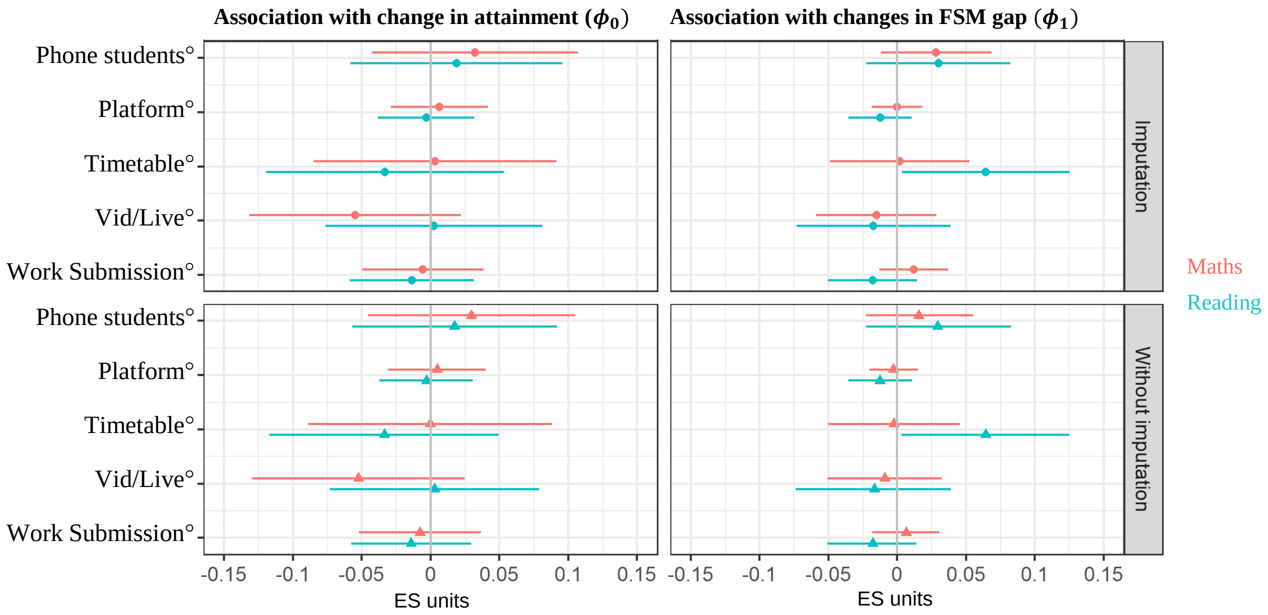
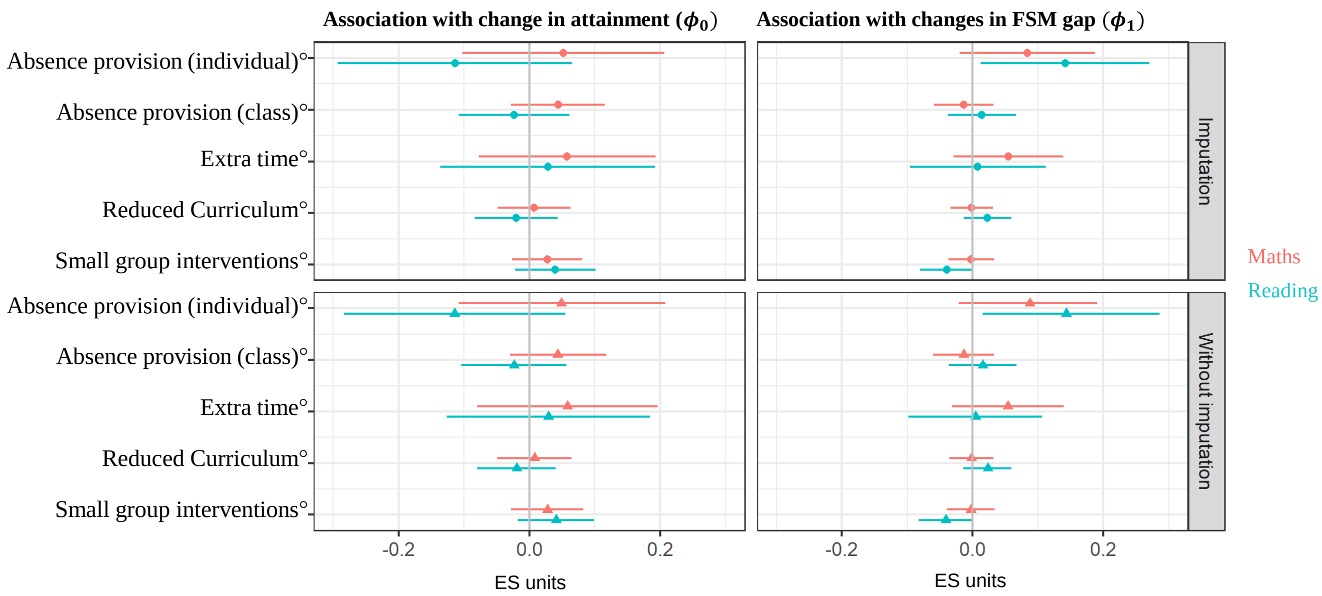


Figure 16: Comparison of imputed and main results for RQ3b



the rule-of-thumb recommended by White *et al.* (2011), and have the same number of imputed datasets as the average percentage of missingness (at student level). This proved to be very time consuming. Given the need to publish this report quickly, we present results that are averaged across 20 imputed data samples. The robustness check is not sensitive to this choice.

Conclusion

Table 16: Summary of key study findings

Research question	Finding
How did the attainment gaps between FSM6 pupils and their peers change between Autumn 2019 and Autumn 2020?	We find evidence that the attainment gaps between FSM eligible pupils and their peers for primary maths have widened since Autumn 2019. On average, for Years 2–6, the maths gap between disadvantaged pupils and their peers widened by an estimated 0.07 ES units (with a plausible range of 0.04 to 0.11). This represents an increase of between 10% to 24% of the pre-covid attainment gap. There was no discernible change in the disadvantage gap for reading.
How did ‘attainment gaps’ change during the Autumn 2020 term, when most pupils returned to ‘face-to-face’ schooling?	There were no clear changes in attainment gaps between FSM eligible pupils and their peers in primary maths or reading, during the Autumn term in 2020. The attainment gap for maths was estimated to change by −0.01 ES units, with a range of [−0.04, 0.02]. For reading the change was 0.01 ES units, with a range of [−0.02, 0.04]. Note that negative numbers indicate a widening of gaps.
What were the associations between how schools responded to ‘remote learning’ from March–May 2020 and changes in attainment gaps between FSM6 pupils and their peers?	We examined five variables that describe primary schools’ responses to remote learning in March–May: ‘phoning students’; ‘timetabling’; ‘live or recorded lessons’; ‘frequency of work submission’; ‘use of technology platforms’. Broadly, we found no evidence of clear associations between these variables and changes in attainment gaps.
What were the associations between school responses during the re-opening period (September–December 2020) and changes in the attainment gaps between FSM6 pupils and their peers in December 2020?	We examined five variables that describe primary school practice in the Autumn 2020 term: ‘providing videos/live streams for absent pupils’; ‘providing videos/live streams for absent classes’; ‘extra learning time’; ‘reducing the curriculum’; and ‘more small group interventions’. Overall, we found limited evidence of clear associations between these variables and changes in attainment gaps. However, there was some tentative evidence that providing video/live lessons to pupils who are absent is associated with reductions in attainment gaps.

Limitations and lessons learned

The study was conceived in the initial lockdown (April 2020) and was set up as quickly as possible. Our goal was to recruit a longitudinal sample in a short period of time, and then administer teacher surveys and tests with a minimal burden on schools.

The study has several important limitations. First, we only have access to a non-random sample of schools and pupils. This does not affect our ability to quantify how inequalities in academic attainment have changed for this set of schools (internal validity)—but there is a risk that a similarly constructed study examining the full population may find different results. We take some comfort from the fact that our analysis samples are similar to broader populations in terms of observable characteristics such as prior attainment (on Key Stage 1) and FSM6 percentages. However, our geographic coverage is not as representative. Moreover, there may be unobserved differences between our sample and the population.

Second, in our analysis of the association between school responses and changes in attainment (RQ3), some of the practices we examined were reported by a small number of schools. We believe that the way in which practices are implemented can have a large impact on their effect. In a study setting where we have no qualitative information about how things were done and we do not have large samples, it is important not to over-interpret null results. Our findings here are tentative.

Third, there were measurement error in the variables capturing school response. Unfortunately, we were unable to link teachers to specific students (thereby linking teacher practice directly with student outcomes). The best we could do was to create school-level variables, despite there being variation within schools. This, again, limits our ability to draw strong conclusions from the analysis in RQ3. In future studies, we strongly recommend that researchers find a way to link teacher responses to pupil outcomes.

Fourth, we used a somewhat novel approach to collecting data for this project. We aimed to minimise administrative burdens on schools by using data from tests that they would have otherwise administered and using existing tools to collect the data. There are disadvantages to this approach. Schools had discretion over when to administer the tests, and with which year groups, classes, and pupils. This limited the numbers of pupils tested at all three periods. While our robustness checks suggest that missingness did not substantially alter our results, if we were repeating a similar exercise in future, we would recommend collecting from schools at the outset of a project a list of which tests they planned to administer (while acknowledging that in difficult circumstances, such as those presented by covid-19, plans would likely change).

Last, we would suggest that studies of future periods of remote learning (or periods of ‘catch-up’) use the survey questions we report here *as a starting point*. Future studies will undoubtedly be interested in aspects of remote or face-to-face learning that differ from the ones we focused on. Even within the confines of this study, we found that shifting from the first set of school closures (March 2020) to the second (in Jan 2021) resulted in inevitable changes to survey questions.

Future research and publications

This project will continue to the end of the 2020–2021 school year. We hope that at least one more round of assessments will be conducted, enabling us to extend the analysis of RQ1.

We have also administered an additional teacher survey, examining the second period of lockdown in January 2021. The survey was sent on March 18th 2021. This was a shortened version of Teacher Survey 1. Our goal is to describe how schools responded to the second period of school closures (Jan–Mar 2021) to support pupils working from home. We hope to use this survey to extend our analysis of RQ3. Covid permitting, these analyses will all be published as an addendum to this report around July 2021.

References

Allen, R., Jerrim, J., and Sims, S. (2020). 'How did the early stages of the covid-19 pandemic affect teacher wellbeing?', UCL Working Paper Working Paper No. 20–15.

Andrew, A., Cattan, S., Costa-Dias, M., Farquharson, C., Kraftman, L., Krutikova, S., Phimister, A., and Sevilla, A. (2020). 'Learning during the lockdown: real-time data on children’s experiences during home learning', available from https//www.ifs.org.uk/publications/14848.

Bacher-Hicks, A., Goodman, J., and Mulhern, C. (2020). *Inequality in household adaptation to schooling shocks: Covid-induced online learning engagement in real time* (0898–2937), National Bureau of Economic Research.

Chetty, R., Friedman, J., Hendren, N., and Stepner, M. (2020). 'The economic impacts of covid-19: Evidence from a new public database built from private sector data', *Opportunity Insights,* available from https//*www.nber.org/papers/w27431.*

Cook, T. D., Shadish, W. R., and Wong, V. C. (2008). 'Three conditions under which experiments and observational studies produce comparable causal estimates: New findings from within-study comparisons', *Journal of Policy Analysis and Management, 27*(4), 724–750.

Cullinane, C., and Montacute, R. (2020). *Covid-19 and social mobility impact brief #1: school closures*. available from https://www.suttontrust.com/wp-content/uploads/2021/01/School-Shutdown-Covid-19.pdf.

Curriculum Associates. (2020). 'Understanding student needs', available from https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/iready-diagnostic-results-understanding-student-needs-paper-2020.pdf.

DfE. (2021). 'Understanding progress in the 2020/21 academic year,' Renaissance Learning, EPI, available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/962330/Learning_Loss_Report_1A_-_FINAL.pdf.

Domingue, B., Hough, H., Lang, D., and Yeatman, D. (2021). *Changing patterns of growth in oral reading fluency during the covid-19 pandemic*. available from https://edpolicyinca.org/publications/changing-patterns-growth-oral-reading-fluency-during-covid-19-pandemic.

Engzell, P., Frey, A., and Verhagen, M. D. (2020). 'Learning inequality during the covid-19 pandemic' available from engzellfreyverhagen_research_nov_2020.pdf (mec.pt).

Hupkau, C., and Petrongolo, B. (2020). 'Work, care and gender during the Covid-19 crisis.' *Fiscal Studies, 41*(3), 623–651.

Juniper Education. (2020). *National Dataset Report*. Juniper Education, available from https://junipereducation.org/wp-content/uploads/juniper_folder/Juniper-Education-National-Benchmark-Dataset-Report.pdf.

Kogan, V., and Lavertu, S. (2021). 'The covid-19 Pandemic and student achievement on Ohio’s third-grade English Language arts assessment', available from http://glenn.osu.edu/educational-governance/reports/reports-attributes/ODE_ThirdGradeELA_KL_1-27-2021.pdf.

Lucas, M., Nelson, J., and Sims, D. (2020a). 'Pupil engagement in remote learning,' NFER, available from https://www.nfer.ac.uk/media/4073/schools_responses_to_covid_19_pupil_engagement_in_remote_learning.pdf.

Lucas, M., Nelson, J., and Sims, D. (2020b). 'Schools' responses to covid-19: Pupil engagement in remote learning,' NFER. available from: https://www.nfer.ac.uk/media/4073/schools_responses_to_covid_19_pupil_engagement_in_remote_learning.pdf.

Maldonado, J., and De Witte, K. (2020). 'The effect of school closures on standardised student test', FEB Research Report Department of Economics.

McCarty, C., and Cooke, C. (2015). *Progress in understanding mathematics assessment*, Hodder Education Group.

McCarty, C., and Ruttle, K. (2016). *PIRA Stage Two (Tests 3–6) Manual: 2ED (Progress in reading assessment)*: Hodder Education Group.

NFER. (2021). 'Impact of school closures and subsequent support strategies on attainment and socio-emotional wellbeing in Key Stage 1: Interim Paper 1,' available from https://educationendowmentfoundation.org.uk/public/files/Publications/Covid-19_Resources/Impact_of_school_closures_KS1_interim_findings_paper_-_Jan_2021.pdf.

Pier, L., Hough, H., Christian, M., Bookman, N., Wilkenfeld, B., and Miller, R. (2021). 'Evidence on learning loss from the CORE data collaborative,' PACE, available from https://edpolicyinca.org/newsroom/covid-19-and-educational-equity-crisis.

Renaissance. (2020). 'How kids are performing,' available from https://www.renaissance.com/how-kids-are-performing/.

RS Assessment. (2020). 'The impact of school closures on autumn 2020 attainment', available from https://www.risingstars-uk.com/rs-assessment/whitepapers

Thomson, D. (2021). 'Do primary schools close the attainment gap for disadvantaged pupils?', FFT Education Datalab, available from https://ffteducationdatalab.org.uk/2021/03/do-primary-schools-close-the-attainment-gap-for-disadvantaged-pupils/.

Weidmann, B., and Miratrix, L. (2020). 'Lurking inferential monsters? Quantifying selection bias in evaluations of school programs', *Journal of Policy Analysis and Management*.

Wong, V. C., Valentine, J. C., and Miller-Bains, K. (2017). 'Empirical performance of covariates in education observational studies', *Journal of Research on Educational Effectiveness, 10*(1), 207–236.

Appendix A: Recruitment documents

Link to email sent to schools to invite them into the project: **https://mailchi.mp/fft/action-required-ffteef-covid-19-research-project?e=%5bUNIQID%5d**

Dear colleague,

I am emailing to request your school's participation in a new national research project investigating **the impact of Covid-19 and school closures on pupil attainment.** The project is being run jointly by FFT and the Education Endowment Foundation (EEF) and your school is of particular importance for the research as you already use FFT Aspire and the standardised tests from RS Assessment (for example, PIRA, PUMA or NTS tests).

Participating in the research project will be straightforward. Over the next 6 months, we will just need your school to:

1. Have some **baseline data** from RS Assessment tests (PIRA/PUMA/NTS) from December 2019 and test these same pupils using **standardised tests** (PIRA/PUMA/NTS) from RS Assessment in September 2020 and December 2020.
2. Ask your teachers and headteacher to complete a **short 10-minute survey** about practices when your school was closed.
3. Give permission for FFT to use your **school's pupil data in FFT Aspire** including the data which we can access from your school MIS and the RS Assessment MARK platform.

Your school will benefit from collaborating with FFT and EEF on this research:

- Each school will **receive a £250 payment** as a thank you for your participation;
- Each school will receive a **'special report' in FFT Aspire** which analyses the impact of covid-19 on the attainment of your pupils compared to other schools.

Your school can choose the number of year groups that participate in this research. You can have one year group participating or multiple year groups.

To register your school for the programme, please click the button below to complete the short online registration form by **15th July**.

Register my school >

If you have any questions or would like any further information about the research project, please do get in touch with Laura James at FFT by emailing: **covid19research@fft.org.uk**.

I hope that you will be able to collaborate with us on this important piece of research. I know that the research will provide a unique and valuable insight into the impact of covid-19 on pupil outcomes for primary schools.

Many thanks,

Paul Charman
Managing Director, FFT

Link to online sign up form: **https://fftedu.wufoo.com/forms/ffteef-covid19-research-project-copy/**

Text from the school sign up form can be seen below.

FFT/EEF covid-19 Research Project Copy

Please read and complete this short form to register your school for the FFT/EEF covid-19 research project.

Please confirm that you will be able to share the following data and information with FFT for the research project: *

☐ ☐ Baseline data: you must have November/December 2019 RS Assessment (PIRA, PUMA, and GAPS or NTS Assessments)

☐ ☐ Outcome data: we will need you to complete two further RS Assessments for these same pupils (September 2020 and November/December 2020)

☐ ☐ Teacher survey: Short survey with headteacher and class teachers about the learning provision for pupils while schools have been closed

☐ ☐ Data collection: Data Exchange connection between FFT Aspire and your School MIS and enabling test data sharing in MARK to FFT Aspire Pupil Tracking

School name *

School DFE number (7 digits) *

Enter a value between **1000000** and **9999999**.

Local Authority *

Barking and Dagenham ▾ Barking and Dagenham ▾

Lead contact within your school for this research project: * First

Last

Email address for the lead contact: *

Lead contact: job role *

◉ ◉ Headteacher / senior leader

○ ○ Middle leader

○ ○ Teacher

○ ○ School administrator

Year groups to be included in the research

For each year group participating in the research we will need:

(1) PIRA/PUMA/GAPS/NTS Assessments test results in MARK;

(2) your teachers to complete a short survey.

Please select the year groups which will be able to participate in the research: *

- ☐ ☐ Current Year 1 (Year 2 in Sept)
- ☐ ☐ Current Year 2 (Year 3 in Sept)
- ☐ ☐ Current Year 3 (Year 4 in Sept)
- ☐ ☐ Current Year 4 (Year 5 in Sept)
- ☐ ☐ Current Year 5 (Year 6 in Sept)

Headteacher and Teacher survey: email addresses

We will need your headteacher and teachers to complete a short 10-minute online survey about teacher practices during the partial school closures.

Headteacher name *First ______ Last ______

Headteacher email address *

Year 1 teachers

Please provide the name and email address for each of your Year 1 teachers

Name ______ First ______ Last

Email address

Name ______ First ______ Last

Email address

Name ______ First ______ Last

Email address

Year 2 teachers

Please provide the name and email address for each of your Year 2 teachers

Name ______ First ______ Last

Email address

Name ______ First ______ Last

Email address

Name ______ First ______ Last

Email address

Year 3 teachers

Please provide the name and email address for each of your Year 3 teachers

Name First Last
Email address
Name First Last
Email address
Name First Last
Email address

Year 4 teachers

Please provide the name and email address for each of your Year 4 teachers

Name First Last
Email address
Name First Last
Email address
Name First Last
Email address

Year 5 teachers

Please provide the name and email address for each of your Year 5 teachers

Name First Last
Email address
Name First Last
Email address
Name First Last
Email address

IMPORTANT: permission to participate in the research project

Name of person completing this form for the school * First Last

Job role of person completing this form *

- Headteacher / senior leader
- Middle leader
- Teacher
- School administrator

By ticking this box, you confirm that you are authorised on behalf of your school to participate in the FFT/EEF covid-19 research project. You confirm that the pupil and teacher information shared with FFT can be used for this research project under the GDPR-compliant data sharing agreement in place between FFT and the school for FFT Aspire. *

- I agree that my school can participate in this research project

Further appendices:

Appendix B: Collecting data

Attainment data
For the purposes of reporting and analysis, schools in the sample used either (or both) MARK (RS Assessment's online analysis tool) or FFT Aspire Pupil Tracking (APT). We collected data from MARK via an interface (API) once schools authorised us to do so. Data was directly extracted from APT, again from schools authorising us to do so. In exceptional circumstances, schools supplied us with some test data in Excel format.

The key variables extracted from the MARK API for each pupil were:

- test suite (PIRA/ PUMA/NTS Reading/ NTS maths);
- national curriculum year (NCY) of the test;
- term of the test (Autumn, Spring, Summer);
- standardised score;
- test date.

Similar variables were also extracted from APT.

Pupil characteristics
Data on pupil characteristics, such as disadvantage, gender, and ethnicity, were collected directly from schools' management information systems (MIS) using Aspire Data Exchange (ADX). Weekly extracts were taken from September 2020 onwards. However, some schools only authorised collection later in the Autumn 2020 term. We did not have the facility to retrospectively collect data.

Published school-level data
Further relevant school characteristics to control for in models, such as inspection ratings, region, historic attainment, and percentage disadvantage were sourced from published government datasets.

Cleaning and transformation

A pseudonymised dataset was constructed from the source data in order to conduct analysis. This involved a number of stages to select, clean, and transform the data to make it suitable for analysis. Stages of data cleaning included:

- identifying relevant test data;
- adding linked pupil characteristic data;
    - determine which FSM6 flag to use;
- equating test scores.

Identifying relevant test data
We first identified relevant test data in both MARK and APT. This meant identifying:

- the type of test (for example, PIRA, PUMA) and the subject;
- the national curriculum year of the test;
- the term of the test;
- the date it was taken.

In some cases, we found that data on the term of test clashed with the date it was taken (for example, Summer tests taken in November). This was particularly the case in data sourced from APT. We adjudicated based on the user-provided test name field.

Where there were multiple results for a pupil for a test, we gave precedence to results that were in the date window we expected, had a UPN, had a non-null score, and had the closest NCY to that of the pupil at the time they took the test.

The process we followed was as follows:

- Identify relevant tests taken by pupils in Years 2 to 5 in Autumn 2019;
  - Date window for tests:
    - MARK: 24/10/2019–14/1/2020
    - APT: October 2019–January 2020
    - Excel: any
- Identify relevant tests taken by pupils in Years 3 to 6 in the 2020–2021 academic year;
  - Date window for Summer tests:
    - MARK: 15/5/2020–31/10/2020
    - APT: June 2020–October 2020
  - Date window for Autumn tests:
    - MARK: 1/11/2020–31/1/2021
    - APT: November 2020–February 2021
- Criteria for determining the year/ subject/ term of test;
  - Subject: test type parsing (for example, PUMA/NTSM tests are classified as maths)
  - Term: MARK and Excel data use the term as provided from source (the term field); APT looks at whether the person providing the data commented on the term they were submitting for in the test_name field or uses the date window otherwise
  - Test NCY: MARK uses the ncyear_id field; APT is a parse of the test_name or falls back on the ncyear_id field (from test source)—these are then compared with the pupil's NCY in school roll data to determine which is more likely.

Linking pupil characteristic data

Test data was linked to data on pupil characteristics using UPN and school identifier (DfE number).

From the weekly snapshots of data collected since the start of Autumn 2020 we:

- identified those on roll at each participating school. this includes those who were on roll at the start of the 2019–2020 academic year, and anyone else who appears in the test data;
- classified any pupils flagged as FSM6 at any point in the Autumn 2020 term as disadvantaged;
- classified pupils according to their SEN status in the first snapshot from the Autumn 2020 term;
- identified other characteristics (ethnicity, EAL, gender, month of birth) based on the modal values across all snapshots.

Appendix C: Teacher survey 1

Table 15: Survey items relating to learning during the first half of the Summer term 2020

1.	How did you set work for your class to complete? Tick the response that best describes how MOST families would usually find out what work is set. • By post or collection from the office • Via a page on the school website • Via an email • Via an online learning or communication platform (for example, Google Classroom, Microsoft Teams, Class Dojo, Seesaw, Firefly etc) • Via a social media site (for example, a school Facebook page) • Via a video or phone call/lesson • Some other way • Not relevant: No work was set for my class	`comms_techplatform` Coded 1 if set or received work via an online learning or communication platform. Coded 0 otherwise, including if no work is set.
2.	How did pupils send completed work back to you? Tick the response that best describes how families would usually send you work. • By post or collection from the office • Via an email • Via an online learning or communication • platform (for example, Google Classroom, Microsoft Teams, Class Dojo, Seesaw, Firefly etc) • Via a social media site (for example, a school Facebook page) • Via a video or phone call/lesson • Some other way • Not relevant: No work was sent in from families or I had no class	

3.	How much did you try to stick to your school's pre-existing curriculum? Tick the response that most closely applies.
	• We tried to stick to the content and pace of our school's curriculum as much as possible
	• We SOMEWHAT reduced the content and/or the pace of the school's curriculum during lockdown
	• We LARGELY PAUSED or ENTIRELY STOPPED the school's curriculum during lockdown and switched mostly to revisiting old topics and retrieval activities
	• We LARGELY PAUSED or ENTIRELY STOPPED the school's curriculum during lockdown and aligned with another curriculum (for example, Oak National or BBC Bitesize topics)
	• None of the above statements align with our practice during lockdown
	• Not relevant / cannot answer
		`curric_div_slowed`
		Coded 1 if teacher some reduced content or slowed pace (0 otherwise).
		
		`curric_div_paused`
		Coded 1 if teacher largely paused or stopped curriculum to align with other curriculum resources (0 otherwise).
		
		`curric_div_retrieval`
		Coded 1 if teacher largely switched to revisiting topics or retrieval practice (0 otherwise).

4.	Which of the following resources have you used as part of the home learning you set for students? Tick all that apply.
	• An app or online SUBSCRIPTION education site (for example, Numbots)
	• An app or online education site WITHOUT subscription needed
	• Videos or streamed content of you teaching
	• Worksheets and tasks that you or your colleagues have created
	• Worksheets and tasks created by someone outside your school
	• BBC Bitesize shows or resources
	• Oak National Academy resources
	• None of the above
	• Not relevant / cannot answer
		`res_site_subs`
		`res_site_free`
		`res_bbc`
		`res_video`
		`res_sheet_ext`
		`res_sheet_int`
		`res_oak`
		Seven binary indicators of whether they used the resource.

5.	Thinking back to the literacy or English resources you provided for home learning, were they originally developed for classroom teaching or for parents to use? Tick the response that most closely applies.	`resource_adapt_fully` Coded 1 if majority of resources and activities were created with parents in mind (0 otherwise).

	• Majority of resources and activities were adapted from classroom teaching activities • Majority of resources and activities were created with parents in mind • Even balance: some were created for parents and some for teachers • Not relevant / cannot answer	`resource_adapt` Coded 1 if either even balance of adaption or fully adapted for parents (0 otherwise).
6.	Thinking back to the literacy or English resources you provided for home learning, could children complete the activities you set without supervision from their parents? Tick the response that most closely applies. • Children could complete all the activities without parental supervision • Children could complete many of the activities without parental supervision • Children could complete some of the activities without parental supervision • Children could not generally complete the activities unless they had parental supervision • Not relevant / cannot answer	`childindependent` Scale of 0–3 where 0 indicates parental supervision needed for everything and 3 indicates all activities designed for working independently.
7.	Did you have any form of daily registration for students learning at home? Tick the response that most closely applies. • Yes, we asked pupils to log onto an online platform or send an email (or similar) every school day • No, but we suggested pupils to log onto an online platform or send an email (or similar) every school day • No daily registration • Not relevant / cannot answer	`interaction_register` Binary indicator coded 1 if a required or suggested daily check-in system was in place.

8.	Did you provide students with a timetable to follow each day (this could be hourly, lesson-by-lesson, or a daily list)?
	• Yes, we asked pupils to follow a daily timetable
	• Yes, but it was only a suggested timetable and pupils were not required to follow it
	• No
	• Not relevant / cannot answer
		`interaction_timetable`
		Binary indicator coded 1 if there was a suggested or required timetable to follow.
		`interaction_timetable_high`
		Binary indicator coded 1 if it was a require timetable to follow.

9.	Were pupils able to do any of the following with their class teacher during the first half of Summer term? (Tick any that apply)
	• Take part in a 'live' online lesson where they could talk
	• Take part in a 'live' online lesson where they could not talk
	• Take part in a 'live' social video chat or check-in time
	• Take part in a text-based online chat where they type conversation
	• Watch a pre-recorded video of their teacher talking
	• None of the above were possible
	• Not relevant / cannot answer
		`interaction_livelesson`
		`interaction_livechat`
		`interaction_textchat`
		`interaction_video`
		
		Four binary indicators of how students could interact with their teacher during lockdown.

10.	Did children (i.e. not the parents) typically speak to you (i.e. their class teacher) on the phone?
	• No, children typically did not speak to their class teacher
	• Yes, one or twice during the half term
	• Yes, about once a week
	• Yes, more than once a week
	• Not relevant / cannot answer
		`interaction_phone`
		Binary indicator for whether student (not parent) spoke to teacher on the phone.
		
		`interaction_phone_regular`
		Binary indicator for whether student spoke weekly to the teacher on the phone.

11.	Did you share examples of student work somewhere for all students to see? Tick any that apply. • Yes: on school website • Yes: in school newsletter sent by email • Yes: in an email to the class (or similar) • Yes: within our online learning platform • Yes: via social media (for example, a Facebook page or Twitter) • Yes: somewhere else not listed above • No • Not relevant / cannot answer	`interaction_worksharing` Binary indicator for whether the school shared examples of student work by any means.
12.	Which best describes the type of feedback you felt able to give on pieces of work submitted by pupils?	`feedback_praise`

	• I didn't give feedback on individual pieces of work • Feedback was essentially all general praise and encouragement • I gave specific feedback intended to support learning on SOME pieces of work • I gave specific feedback intended to support learning on MANY pieces of work • I gave specific feedback intended to support learning on all/ALMOST ALL pieces of work • Not relevant / cannot answer	Binary indicator for whether feedback given was mostly praise and encouragement. `feedback_specific` Binary indicator for whether teacher gave any specific feedback on work intended to support learning.

13.	How frequently did you suggest that parents or pupils send in work?
	• Every day (or after each lesson)
	• Several times a week
	• Once a week
	• Once a fortnight
	• Less than once a fortnight
	• I gave no suggestion about how frequently work should be sent in
	• Not relevant / cannot answer
		`feedback_given`
		Binary indicator for whether or not the teacher encouraged work to regularly be submitted.
		
		`feedback_given_daily`
		Binary indicator for whether or not the teacher encouraged the daily submission of work.

14.	How much do you agree with the following statement: 'During lockdown whilst most of my students were learning at home, it was easy for me to monitor who was, and wasn't, completing work.' • Strongly agree • Somewhat agree • Slightly agree • Slightly disagree • Somewhat disagree • Strongly disagree • Cannot answer / not relevant	`monitoring_ease` Scale from 1–6 where 6 indicates high ease in monitoring students.
15.	How much do you agree with the following statement: 'When setting work for remote learning, I found it difficult to differentiate to the lowest attainers in my class.' • Strongly agree • Somewhat agree • Slightly agree • Slightly disagree • Somewhat disagree • Strongly disagree • Cannot answer / not relevant	`differentiate_ease` Scale from 1–6 where 6 indicates that the teacher found differentiation a challenge during lockdown.
16.	Overall, how good do you feel the home learning experience was for your class? Tick the response that most closely aligns with your feelings. • Very successful: all or almost all pupils were consistently completing school work during the first half of Summer term • Successful: the majority of pupils seemed to be completing school work • Mixed: whilst many pupils did seem to be completing school work, a significant portion of the class were not doing so much • Not so good overall: most pupils were clearly completing far less work than we had set for them • Other…	`Outcome` Coded from 0 to 3 but not used in analysis.

Notes: (1) Cannot answer / not relevant responses generally coded as missing
(2) Preceding these questions, information on the teacher's school, class, and job role were collected.

Table 16: Rotated factor loadings (pattern matrix): variables created from survey 1

Variable	Factor1 λ = 2.95	Factor2 λ = 1.29	Factor3 λ = 1.16	Factor4 λ = 1.12	Factor5 λ = 0.97	Factor6 λ = 0.81	Factor7 λ = 0.61	Factor8 λ = 0.45
comms_techplatform	−0.03	**0.27**	−0.01	0.02	−0.08	0.02	−0.02	−0.16
curric_div_retrieval	0.00	0.05	**0.37**	0.02	0.07	−0.09	0.00	0.01

res_site_subs	0.00	−0.01	0.05	0.02	−0.03	0.08	0.04	**0.16**
res_site_free	0.02	−0.02	−0.04	0.03	−0.03	**0.19**	0.04	0.03
res_bbc	0.00	0.01	0.01	−0.02	0.00	**0.36**	−0.02	0.03
res_video	−0.01	−0.03	0.00	−0.04	0.06	0.03	**0.35**	0.04
res_sheet_ext	−0.01	0.01	0.00	0.01	−0.04	**0.18**	0.13	0.00
res_sheet_int	−0.02	0.03	−0.02	0.00	−0.01	0.02	0.10	0.00
res_oak	0.02	0.00	0.02	0.01	0.07	**0.22**	−0.10	−0.13
resource_adapt	0.00	0.02	0.01	**0.44**	0.01	0.03	0.03	0.00
resource_adapt_fully	0.01	0.03	−0.03	**0.44**	0.01	−0.01	−0.03	−0.03
childindependent	0.00	0.09	0.02	−0.06	0.02	0.02	−0.07	−0.07
interaction_register	0.02	**0.11**	−0.01	0.04	0.10	0.01	−0.05	0.04
interaction_timetable	0.00	0.02	−0.03	0.03	0.03	0.03	0.03	**0.24**
interaction_timetable_high	0.00	0.07	−0.06	0.00	0.07	−0.02	−0.07	**0.22**
interaction_livelesson	−0.02	0.05	−0.02	−0.03	**0.16**	0.03	−0.06	0.03
interaction_livechat	−0.02	**0.07**	0.02	−0.02	0.04	0.10	0.04	0.04
interaction_textchat	0.00	**0.11**	0.04	−0.02	−0.03	0.00	0.00	−0.07
interaction_video	0.01	0.01	0.00	0.03	−0.03	−0.06	**0.34**	−0.05
interaction_phone	0.00	−0.03	0.00	0.00	**0.35**	0.07	0.00	−0.09
interaction_phone_regular	0.01	−0.04	0.03	0.02	**0.33**	−0.08	0.03	0.07
interaction_worksharing	−0.02	0.01	0.00	−0.03	**0.11**	0.02	0.06	0.01
feedback_specific	**0.52**	0.09	0.02	−0.01	−0.01	−0.14	0.14	−0.20
feedback_praise	**−0.46**	0.27	−0.02	−0.04	0.05	−0.09	0.15	−0.14
feedback_given	0.01	**0.19**	−0.01	0.00	−0.03	−0.06	−0.05	0.20
feedback_given_daily	−0.02	**0.23**	0.03	0.01	−0.02	0.04	−0.02	0.12
monitoring_ease	−0.02	**0.25**	0.01	0.01	0.02	0.02	0.06	−0.11
differentiate_ease	0.01	−0.02	0.04	0.00	−0.10	0.02	0.01	**0.14**
curric_slowdown	0.02	0.00	**0.50**	−0.03	−0.02	0.04	0.01	0.04

Appendix D: Teacher survey 2

Table 17: Mapping of teacher survey 2 questions to variables used for analysis

Question	Responses	Variable
1. Which of the following statements are true about your class and school absence so far since September?	2 = no class isolation needed within school; 1 = class isolation with live lessons; 0 = class isolation without live lessons	Absence provision for covid
2. Which of the following types of provision were available for an individual child who needed to isolate whilst their peers were in class?	0 = no access to any video/streaming of class; 1 = video/streaming of class	Class provision for covid
3. Has your school done any of the following to compensate for lost learning time last year?		
3a. Removing items from the curriculum	0 = no; 1 = yes	Reduced curriculum
3b. Running extra lunchtime learning activities	0 = no; 1 = yes	Extra time
3c. Extending the school day	0 = no; 1 = yes	Extra time
3d. Being open during holidays	0 = no; 1 = yes	Extra time
3e. Small group interventions DURING lessons	0 = no; 1 = yes	Small group interventions
3f. Small group interventions OUTSIDE lesson time	0 = no; 1 = yes	Small group interventions
4. How many extra minutes a day (over and above a normal year) were you spending on English and maths last term to cope with missed learning as a result of shutdown and absences? Tick the response that most closely applies.	0 = no extra time; 1 = 15 mins a day; 2 = 30 mins a day; 3 = 45 mins a day or more	Extra time
5. Was your school making greater or less use of small group and individual face-to-face interventions (for example, run by TAs) last term?	0 = fewer interventions; 1 = same number as normal; 2 = more interventions	Small group interventions

Notes: we asked two more questions in addition to those presented here. At the time of writing, we did not have access to responses to the following question ‘Last term, how much time each day have you typically spent supporting remote learning for isolating students, whilst also teaching a class?’. We hope to include this in the final draft. The final question was ‘In which of these subjects have you returned to (re-)teaching topics and skills that the class had missed as a result of lockdown during Summer term 2020? Tick all that apply’. We excluded this on the grounds that estimated ICC was so low that analysis at the school level did not seem worthwhile.

Table 18: Correlation of items from teacher survey 2

Variable description		q1*	q2	q3a	q3b	q3c	q3d	q3e	q3f	q4
class_absence provision	q2	0.06								
curriculum (remove items)	q3a	−0.06	−0.12							
extra time (lunch)	q3b	0.02	−0.12	0.1						
extra time (school day)	q3c	0.01	0.09	0.02	0.01					
extra time (holidays)	q3d	0.13	0.01	0.08	0.1	0.16				
small group (during lessons)	q3e	−0.15	0.08	0.15	0.07	0.08	0.14			
small group (outside lesson time)	q3f	0.2	−0.16	0	0.17	−0.06	0.04	0.27		
extra English/maths	q4	0.01	0.12	0.44	−0.01	0.18	0.15	0.06	−0.09	
extra small group interventions	q5	0.16	0.21	−0.11	0.18	−0.13	0.1	0.45	0.44	−0.11

Notes: q1* is ‘pupil_absence provision

Appendix E: Likelihood ratio tests for research question 1

In this appendix, we compare two sets of models:

1. Models where we estimate Δ_{g}^{k} separately for each year group, for example, having the following structure:[53]

$$Y_{ijT_1}^{k} - Y_{ijT_0}^{k} = \alpha_j + \delta_{T_1g}^{k} + \Delta_{T_1g}^{\mathrm{k}} F_i + \mathrm{month}_{iT_0} + \beta\mathrm{month\_difference}_i + e_{ijg}^{k}$$

$$\alpha_j \sim N(0, \sigma_\alpha^2)$$

$$e_{ijg}^{k} \sim N(0, \sigma^2)$$

2. Models where we constrain Δ^k parameter to be constant across year levels (to provide an 'average' change in disadvantage gaps), for example, having the following structure:[54]

$$Y_{ijT_1}^{k} - Y_{ijT_0}^{k} = \alpha_j + \delta_{T_1g}^{k} + \Delta_{T_1}^{\mathrm{k}} F_i + \mathrm{month}_{iT_0} + \beta\mathrm{month\_difference}_i + e_{ijg}^{k}$$

$$\alpha_j \sim N(0, \sigma_\alpha^2)$$

$$e_{ijg}^{k} \sim N(0, \sigma^2)$$

The output comparing these models using likelihood ratio tests is below.

Table 19: LR tests for reading

	Autumn 2019–Summer 2020	Summer 2020–Autumn 2020	Autumn 2019–Autumn 2020
Fixed parameters in model 1	13	13	13
AIC	19,230.630	17,298.550	19,500.660
BIC	19,322.030	17,389.950	19,592.060
logLik	−9,602.315	−8,636.274	−9,737.329
deviance	19,204.630	17,272.550	19,474.660
Chisq	3.094	1.325	0.615
Df	4	4	4
Pr(> Chisq)	**0.542**	**0.857**	**0.961**

Table 20: LR tests for maths

	Autumn 2019–Summer 2020	Summer 2020–Autumn 2020	Autumn 2019–Autumn 2020
Fixed parameters in model 1	13	13	13
AIC	17,586.220	15,221.050	17,011.510
BIC	17,677.890	15,312.720	17,103.190
logLik	−8,780.109	−7,597.525	−8,492.757
deviance	17,560.220	15,195.050	16,985.510
Chisq	5.564	5.164	7.330
Df	4	4	4
Pr(> Chisq)	**0.234**	**0.271**	**0.119**

[53] This example is for the period Autumn 2019–Summer 2020.
[54] This example is for the period Autumn 2019–Summer 2020.

Appendix F: Regression results for research question 1

Table 21: Regression results for research question 1

	Maths			*Reading*		
	Autumn 2019–Summer 2020	Summer 2020–Autumn 2020	Autumn 2019–Autumn 2020	Autumn 2019–Summer 2020	Summer 2020–Autumn 2020	Autumn 2019–Autumn 2020
	(1)	(2)	(3)	(4)	(5)	(6)
Autumn 2019 test date: November 2019	−0.017		−0.074	0.028		0.079
	(0.040)		(0.042)	(0.044)		(0.048)
Autumn 2019 test date: October 2019 or earlier	0.040		0.010	0.329**		0.238*
	(0.119)		(0.115)	(0.108)		(0.120)
Summer 2020 test: September 2020		0.141*			0.248**	
		(0.064)			(0.090)	
Summer 2020 test: August 2020 or earlier		−0.120			0.312**	
		(0.084)			(0.109)	
Difference in test dates	−0.015	0.072**	0.010	−0.018	−0.054**	−0.003
	(0.010)	(0.015)	(0.027)	(0.012)	(0.017)	(0.030)
Year 3	0.115**	−0.091**	0.046*	0.114**	−0.091**	0.034
	(0.025)	(0.022)	(0.023)	(0.029)	(0.026)	(0.028)
Year 4	0.063**	0.051*	0.113**	0.086**	0.034	0.123**
	(0.024)	(0.021)	(0.023)	(0.027)	(0.024)	(0.027)
Year 5	0.032	0.033	0.068**	0.046	−0.004	0.046
	(0.024)	(0.021)	(0.023)	(0.027)	(0.024)	(0.028)
Year 6	0.036	0.026	0.057*	0.030	0.038	0.077**
	(0.024)	(0.021)	(0.024)	(0.028)	(0.025)	(0.028)
FSM6	−0.065**	−0.008	−0.075**	−0.010	0.010	−0.003
	(0.017)	(0.015)	(0.017)	(0.019)	(0.017)	(0.020)
Constant	0.171	−0.303**	−0.089	0.146	−0.113	−0.023
	(0.092)	(0.068)	(0.314)	(0.104)	(0.092)	(0.353)
Observations	8,534	8,534	8,534	8,359	8,359	8,359

Key: $*p < 0.05$, $**p < 0.01$

Table 22: MATHS results by year level for research question 1

	Change in score, Autumn 2019–Summer 2020					Change in score, Summer 2020–Autumn 2020					Change in score, Autumn 2019–Autumn 2020				
	(Year 2)	(Year 3)	(Year 4)	(Year 5)	(Year 6)	(Year 2)	(Year 3)	(Year 4)	(Year 5)	(Year 6)	(Year 2)	(Year 3)	(Year 4)	(Year 5)	(Year 6)
Autumn 2019 test date: November 2019	0.434**	−0.211	−0.051	0.177*	0.150						0.245	−0.075	−0.121	−0.064	0.061
	(0.134)	(0.172)	(0.074)	(0.075)	(0.098)						(0.166)	(0.125)	(0.088)	(0.094)	(0.097)
Autumn 2019 test date: October 2019 or earlier	0.231	−0.006	0.270	0.192	−0.175						−0.128	0.563	0.148	0.320	−0.097
	(0.338)	(0.358)	(0.150)	(0.173)	(0.169)						(0.467)	(0.352)	(0.197)	(0.229)	(0.184)
Summer 2020 test date: September 20 20						0.470	0.112	0.389	0.275	0.291					
						(0.255)	(0.142)	(0.204)	(0.173)	(0.156)					
Summer 2020 test date: August 2020 or earlier						0.449	0.917	0.472*	0.187	−0.004					
						(0.312)	(0.540)	(0.240)	(0.227)	(0.182)					
Difference in test dates	0.034	−0.031	0.040	−0.023	−0.003	0.057	0.041	−0.013	−0.002	0.098**	0.105	−0.280**	0.050	0.136	−0.036
	(0.064)	(0.102)	(0.026)	(0.022)	(0.022)	(0.063)	(0.076)	(0.038)	(0.041)	(0.030)	(0.139)	(0.091)	(0.060)	(0.078)	(0.061)
FSM6	−0.087	−0.083	−0.034	−0.029	−0.073*	−0.027	−0.051	−0.003	0.020	0.028	−0.123**	−0.135**	−0.040	−0.011	−0.050
	(0.045)	(0.047)	(0.032)	(0.030)	(0.031)	(0.040)	(0.039)	(0.029)	(0.027)	(0.029)	(0.045)	(0.042)	(0.031)	(0.029)	(0.029)
Constant	−0.463	0.476	−0.262	0.185	0.082	−0.659*	−0.320	−0.368	−0.204	−0.521**	−1.372	3.402**	−0.494	−1.583	0.480
	(0.549)	(0.914)	(0.221)	(0.188)	(0.187)	(0.273)	(0.173)	(0.205)	(0.178)	(0.160)	(1.635)	(1.074)	(0.708)	(0.922)	(0.710)
Observations	1,295	1,261	1,950	2,273	1,755	1,295	1,261	1,950	2,273	1,755	1,295	1,261	1,950	2,273	1,755

Key: *$p < 0.05$, **$p < 0.01$

Table 23: READING results by year level for research question 1

	Change in score, Autumn 2019–Summer 2020					Change in score, Summer 2020–Autumn 2020					Change in score, Autumn 2019–Autumn 2020				
	(Year 2)	(Year 3)	(Year 4)	(Year 5)	(Year 6)	(Year 2)	(Year 3)	(Year 4)	(Year 5)	(Year 6)	(Year 2)	(Year 3)	(Year 4)	(Year 5)	(Year 6)
Autumn 2019 test date: November 2019	0.128	0.087	0.125	0.129	0.222*						−0.054	0.087	0.026	0.200	0.145
	(0.166)	(0.206)	(0.089)	(0.104)	(0.090)						(0.186)	(0.152)	(0.107)	(0.121)	(0.099)
Autumn 2019 test date: October 2019 or earlier	0.284	0.115	0.298	0.383	0.085						−0.013	0.252	0.485*	0.758**	0.094
	(0.360)	(0.404)	(0.163)	(0.215)	(0.180)						(0.473)	(0.355)	(0.218)	(0.275)	(0.225)
Summer 2020 test date: September 2020						−0.050	0.245	0.450*	0.431*	0.438					
						(0.192)	(0.163)	(0.223)	(0.190)	(0.241)					
Summer 2020 test date: August 2020 or earlier						−0.190	−0.644	0.506	0.536*	0.206					
						(0.280)	(0.635)	(0.263)	(0.225)	(0.275)					
Difference in test dates	0.006	−0.050	0.044	0.043	−0.013	0.012	−0.083	−0.045	−0.049	0.037	0.145	−0.092	−0.081	−0.093	0.021
	(0.054)	(0.146)	(0.029)	(0.027)	(0.025)	(0.062)	(0.089)	(0.041)	(0.039)	(0.038)	(0.149)	(0.125)	(0.064)	(0.078)	(0.061)
FSM6	−0.054	0.018	−0.00004	−0.033	−0.026	0.019	0.004	0.011	0.020	0.003	−0.039	0.018	0.003	−0.018	−0.024
	(0.050)	(0.049)	(0.036)	(0.036)	(0.038)	(0.045)	(0.041)	(0.035)	(0.032)	(0.037)	(0.050)	(0.049)	(0.038)	(0.038)	(0.037)
Constant	−0.094	0.538	−0.375	−0.380	0.061	0.052	−0.138	−0.334	−0.333	−0.429	−1.711	1.032	0.987	1.027	−0.253
	(0.465)	(1.314)	(0.246)	(0.231)	(0.212)	(0.194)	(0.221)	(0.229)	(0.200)	(0.236)	(1.748)	(1.473)	(0.753)	(0.919)	(0.719)
Observations	1,295	1,217	1,911	2,190	1,746	1,295	1,217	1,911	2,190	1,746	1,295	1,217	1,911	2,190	1,746

Key: *$p < 0.05$, **$p < 0.01$

Appendix G: Regression results for research question 3

Table 24: Regression results for research question 3a

	Change in maths score, Autumn 2019–Summer 2020					Change in reading score, Autumn 2019–Summer 2020				
	Platform use	Time-tabling	Work sub-missions	Phone calls	Vid/live lessons	Platform use	Time-tabling	Work sub-missions	Phone calls	Vid/live lessons
FSM6	−0.046**	−0.045**	−0.046**	−0.046**	−0.045**	−0.012	−0.009	−0.010	−0.011	−0.011
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.018)	(0.018)	(0.019)	(0.019)	(0.019)
% pupils achieving expected standard in R/W/M in 2019	0.004	0.003	0.011	0.010	0.020	0.067	0.053	0.078	0.075	0.063
	(0.228)	(0.228)	(0.229)	(0.227)	(0.225)	(0.224)	(0.223)	(0.223)	(0.224)	(0.224)
% pupils achieving higher standard R/W/M in 2019	−0.059	−0.078	−0.099	−0.024	−0.064	−0.193	−0.184	−0.241	−0.157	−0.173
	(0.478)	(0.473)	(0.476)	(0.473)	(0.466)	(0.483)	(0.476)	(0.486)	(0.477)	(0.476)
Ofsted 2	−0.041	−0.042	−0.041	−0.043	−0.036	0.084	0.084	0.081	0.091	0.086
	(0.104)	(0.104)	(0.104)	(0.103)	(0.102)	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)
Ofsted 3	−0.044	−0.040	−0.036	−0.044	−0.035	0.050	0.054	0.050	0.046	0.046
	(0.121)	(0.120)	(0.120)	(0.118)	(0.117)	(0.121)	(0.120)	(0.120)	(0.120)	(0.120)
Ofsted 4	0.043	0.044	0.045	0.028	0.038	0.200	0.194	0.187	0.197	0.202
	(0.172)	(0.172)	(0.172)	(0.172)	(0.170)	(0.171)	(0.171)	(0.171)	(0.171)	(0.171)
% pupils who were FSM6 in 2019	−0.320	−0.334	−0.333	−0.334	−0.314	−0.369*	−0.366*	−0.360*	−0.357*	−0.356*
	(0.185)	(0.182)	(0.178)	(0.177)	(0.176)	(0.187)	(0.182)	(0.179)	(0.179)	(0.179)
Female	−0.001	−0.001	−0.001	−0.001	−0.001	0.027*	0.027*	0.027*	0.027*	0.027*
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
EAL	0.004	0.003	0.011	0.010	0.020	0.067	0.053	0.078	0.075	0.063
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Main effect of school response variable (Z)	0.005	−0.001	−0.008	0.030	−0.053	−0.003	−0.034	−0.014	0.017	0.003
	(0.018)	(0.045)	(0.022)	(0.038)	(0.039)	(0.017)	(0.042)	(0.022)	(0.037)	(0.038)
Interaction of school response and FSM (Z:FSM6)	−0.003	−0.002	0.007	0.015	−0.009	−0.013	0.064*	−0.018	0.029	−0.016
	(0.009)	(0.024)	(0.012)	(0.019)	(0.021)	(0.012)	(0.031)	(0.016)	(0.026)	(0.028)
Constant	0.244	0.251	0.246	0.243	0.229	0.057	0.064	0.054	0.040	0.052
	(0.202)	(0.203)	(0.201)	(0.201)	(0.199)	(0.204)	(0.203)	(0.201)	(0.202)	(0.202)
Observations	13,261	13,261	13,261	13,261	13,261	12,564	12,564	12,564	12,564	12,564

Key: **p* < 0.05, ***p* < 0.01

Table 25: Regression results for research question 3b

	Change in maths score, Summer 2020– Autumn 2020					Change in reading score, Summer 2020– Autumn 2020				
	Pupil absence	School closures	Reduce curriculum	Extra time	Small group	Pupil absence	School closures	Reduce curriculum	Extra time	Small group
FSM6	−0.031*	−0.031	−0.030	−0.029	−0.030	−0.006	−0.001	−0.004	−0.003	−0.005
	(0.016)	(0.017)	(0.017)	(0.016)	(0.017)	(0.019)	(0.019)	(0.018)	(0.019)	(0.019)
% pupils achieving expected standard in R/W/M in 2019	−0.329	−0.357	−0.363	−0.349	−0.345	0.033	0.078	0.080	0.081	0.111
	(0.210)	(0.206)	(0.208)	(0.207)	(0.207)	(0.223)	(0.220)	(0.221)	(0.221)	(0.219)
% pupils achieving higher standard R/W/M in 2019	1.129**	1.050*	1.109*	1.018*	1.115**	0.778	0.860	0.797	0.822	0.853
	(0.430)	(0.427)	(0.436)	(0.430)	(0.425)	(0.457)	(0.464)	(0.470)	(0.463)	(0.456)
Ofsted 2	−0.252*	−0.297*	−0.274*	−0.245*	−0.241*	−0.225	−0.175	−0.175	−0.181	−0.174
	(0.119)	(0.117)	(0.117)	(0.118)	(0.119)	(0.146)	(0.151)	(0.149)	(0.150)	(0.148)
Ofsted 3	−0.226	−0.274*	−0.255	−0.218	−0.229	−0.222	−0.170	−0.173	−0.171	−0.172
	(0.136)	(0.133)	(0.133)	(0.136)	(0.134)	(0.163)	(0.166)	(0.165)	(0.168)	(0.163)
Ofsted 4	−0.299	−0.341	−0.330	−0.320	−0.304	−0.229	−0.175	−0.179	−0.179	−0.172
	(0.183)	(0.177)	(0.180)	(0.178)	(0.180)	(0.207)	(0.209)	(0.210)	(0.209)	(0.206)
% pupils who were FSM6 in 2019	0.375*	0.443*	0.367*	0.389*	0.416*	0.219	0.175	0.221	0.217	0.282
	(0.170)	(0.181)	(0.171)	(0.170)	(0.176)	(0.180)	(0.196)	(0.182)	(0.182)	(0.185)
Female	−0.017	−0.017	−0.017	−0.017	−0.017	−0.026	−0.026	−0.027	−0.026	−0.026
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
EAL	0.028	0.029	0.028	0.027	0.029	−0.005	−0.005	−0.004	−0.004	−0.001
	(0.019)	(0.020)	(0.020)	(0.020)	(0.020)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Main effect of school response variable ($\tilde{Z}$)	0.049	0.043	0.008	0.058	0.027	−0.115	−0.024	−0.020	0.029	0.041
	(0.079)	(0.037)	(0.028)	(0.069)	(0.028)	(0.085)	(0.040)	(0.030)	(0.078)	(0.029)
Interaction of school response and FSM ($\tilde{Z}$:FSM6)	0.087	−0.013	−0.002	0.054	−0.003	0.143*	0.016	0.023	0.005	−0.041*
	(0.052)	(0.023)	(0.017)	(0.043)	(0.018)	(0.065)	(0.026)	(0.018)	(0.052)	(0.021)
Constant	0.208	0.250	0.255	0.225	0.199	−0.027	−0.096	−0.106	−0.105	−0.152
	(0.206)	(0.196)	(0.198)	(0.198)	(0.204)	(0.234)	(0.232)	(0.233)	(0.234)	(0.236)
Observations	7,743	7,743	7,743	7,743	7,743	7,312	7,312	7,312	7,312	7,312

Key: *$p < 0.05$, **$p < 0.01$

Appendix H: Regression results for missingness analysis

Research question 1

	Dependent variable: full responder for RQ1 (binary variable)	
	Maths	Reading
Autumn 2019 score	0.109**	0.072**
	(0.016)	(0.016)
FSM6	−0.100**	0.020
	(0.033)	(0.033)
EAL	0.295**	0.437**
	(0.037)	(0.037)
SEN	−0.025	−0.068
	(0.043)	(0.044)
Female	0.016	−0.001
	(0.030)	(0.030)
London school	−1.810**	−1.807**
	(0.057)	(0.051)
Northern school	−0.601**	−0.469**
	(0.035)	(0.036)
Constant	0.010	−0.098**
	(0.028)	(0.027)
Observations	20,053	20,473
Notes:		$^{*}p < 0.05$, $^{**}p < 0.01$

Research question 3

	Dependent variable: full responder for RQ3			
	Maths RQ3a	Reading RQ3a	Maths RQ3b	Reading RQ3b
Autumn 2019 score	1.369**	0.765**		
	(0.138)	(0.158)		
Summer 2020 score			−0.077	0.245**
			(0.290)	(0.038)
FSM6	−1.400**	−1.337**	−1.259*	0.046
	(0.250)	(0.322)	(0.559)	(0.076)
EAL	−0.097	−0.268	0.357	0.534**
	(0.251)	(0.323)	(0.653)	(0.087)
SEN	−0.046	0.135	1.198	−0.149
	(0.247)	(0.363)	(1.076)	(0.092)
Female	0.496*	0.442	0.310	0.150*
	(0.232)	(0.315)	(0.548)	(0.070)
Constant	6.685**	6.511**	6.515**	1.795**
	(0.299)	(0.329)	(0.518)	(0.060)
Observations	13,481	12,741	7,893	7,893

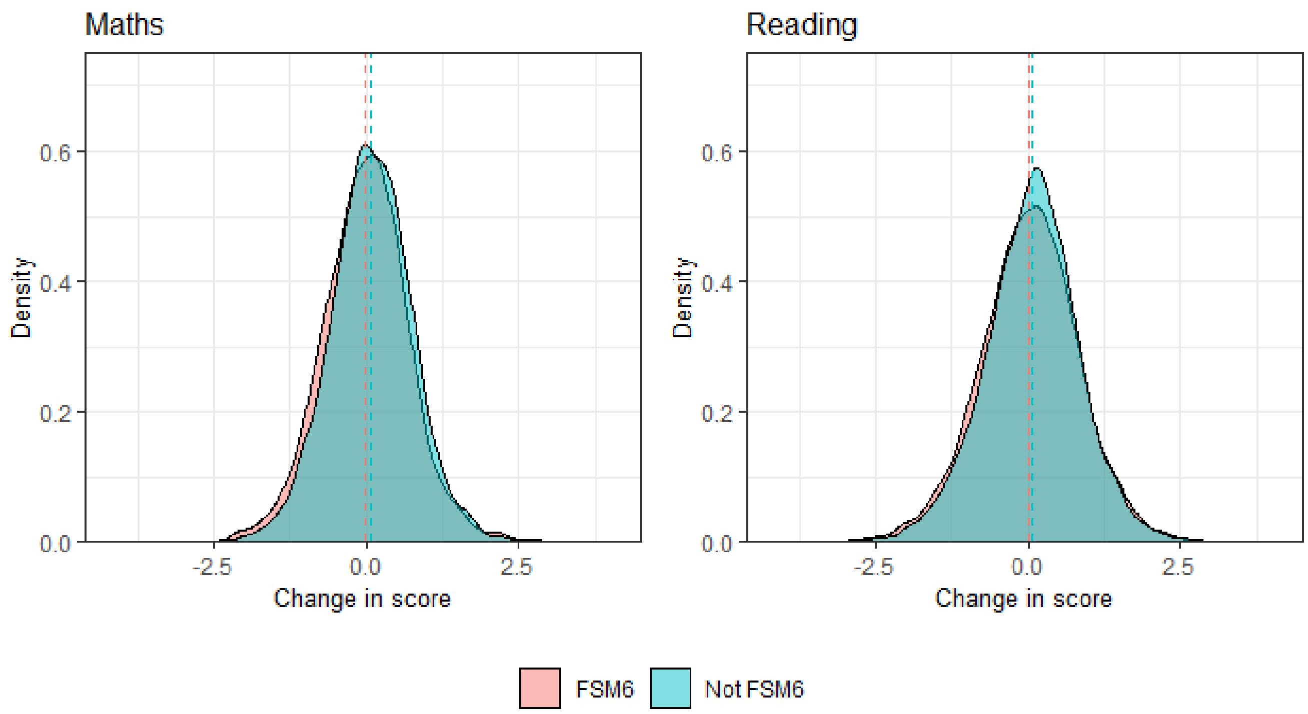
$^{**}p < 0.05$, $^{**}p < 0.01$

Note: we did not include London and northern indicators, as these made models unstable.

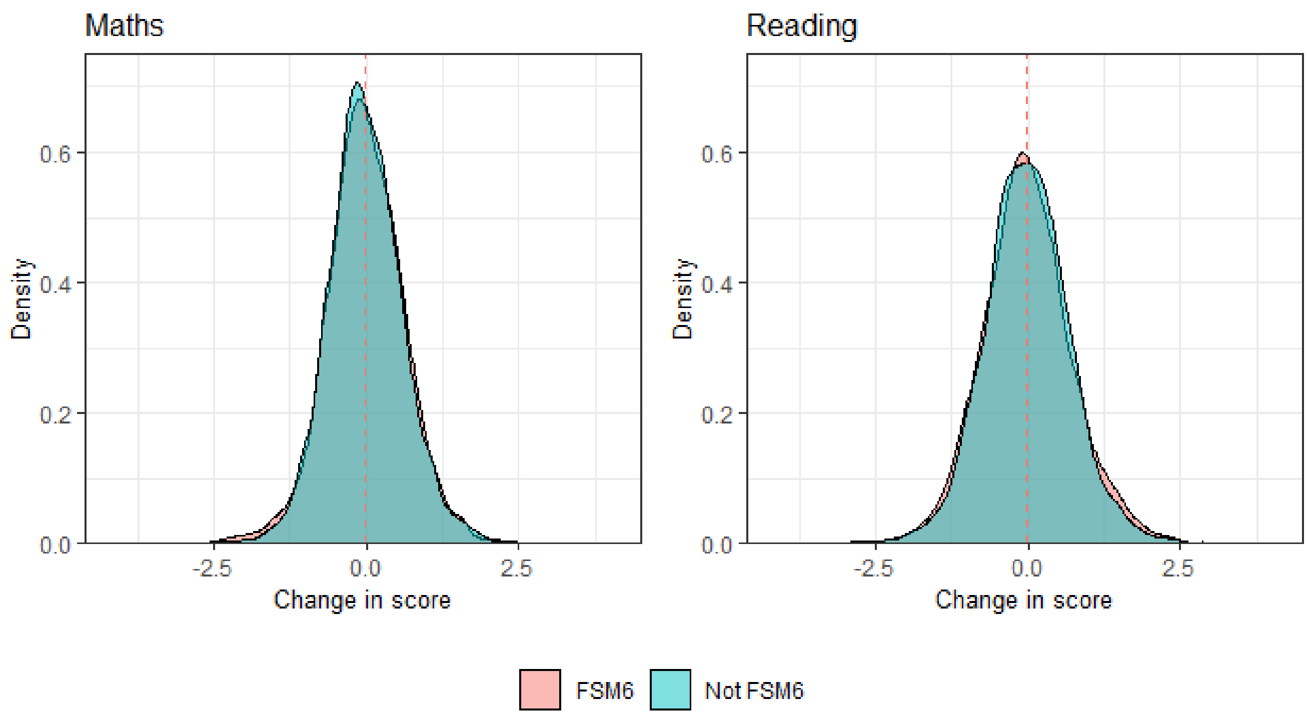
Appendix I: Distribution of change in attainment

This appendix shows the distribution of change in attainment for two periods (relevant to research questions 3a and 3b).

RQ3a: distribution of change in attainment from Autumn 2019 to Summer 2020



RQ3b: distribution of change in attainment from Summer 2020 to Autumn 2020



Education Endowment Foundation

The Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
London
SW1P 4QP

https://educationendowmentfoundation.org.uk

@EducEndowFoundn

Facebook.com/EducEndowFoundn